

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGCGGTCTGGAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCCCTCAGCCCTGTAATTGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTATGCATTG
CTACCATTATGTCGTATAAGCAAGTTCATGCTTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTGCTGCACATGTAAGTGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTCAGACCATCCTTCTACCAAATGCAGCCAAAATCCATGGCAAACAAAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC
ATCAGTTTGCACTGGCAATTGGGACTGATTAGAACAGAAACTCCATTGGAACCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAACGGTCTATGTCATTTCCTTCTT
GGTTTTCTGACTTACATTGTAAGTTCAGAAAATTCTTACGGGTGGAAGCCAATTAC
TGGATTAACCCCTATGACACTGCACCTGCCCTATTAACAATGAACGAACACGGCTACTTCCA
GAGATATTTGAAGGATAAAATATTCTGTAATGATTATGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSAWSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAEWSMSFSFFGFLTYIRDFOKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACCGCTGGCGGACCGTGGGGAGAGCCGCAGTCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGTGTGGCCTCGGGGAGTGGAAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCTGAGTTCCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTTCTCATGCCCAATTGTTAAAGACTATGAGATACTGAGTCAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTGCATTTCTGACCATGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCGTTATTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG
TTTCATGGTGCCTTTTACATTGGCTATTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
CTGCTTTCTGTCTCTTATGGCTGACCTTATGTATTCTCTGGAAACTAGGAGATCCCTTCC
CATTCTCAGCCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTGGATTTGGTGTGCAACTGCCCATACACTTACATGTCTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCAGTGTGCAAACCATGGATATGAT
CATAAAGAAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCACTTCAGCATCAGGAAGTGAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGAAAGATAAGCAGGCAGCTTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAACCTCAAGGGAAATATTAAATTCTGGTT
ACTTTCTCTATTACTGTGTTGGAAATTTCATGGCTACCATCAATATTGTTTGATCGAGTT
GGGAAACGGATCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCCACACATTCTTCATTCTGTTGAATAATCATCGTCACATCCATCAGGAGAT
TGCTGATCACTCTTACCAAGTTCTTATGGCATCTCTAGCAGTAAGTCCCAATGTCATTGCT
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGTCATGGAAATGAGTATGCCTT
AGAATACCGCACCATAATCACTGAAGTCCTGGAGAACACTGCAGTTCAACTTCTATCACC
GGTTGGTGTGACCAAGGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATGTGATCTTCTGGTCAGCGCTCTCTAGCATACTCTCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAAATT
GATATAAGAGGGGGAAAAATGGAACCAGGGCTGACATTAAACAAACAAAATGCTATGGTAGC
ATTTCACCTCATGCAACTCCTCCCGTCAGGTGATACTATGACCATGAGTAGCATGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGGAGCATCCGTGTGGATATGAGGCTGG
TGTAGAGGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
CTATGGTAGCTGAGCCAAACACGTAGGATTCCGTTAAGGTTACATGGAAAGGTTAGCTT
CCTGAGATTGACTCATTAAACAGAGACTGTAACAAAAAAAAGGGCGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCATGCCAACTTGTATTGAGCTTATAATG

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV

FIGURE 4

LNSSSRYFHWKMNLCVILLILVFMPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGD
FPILSPKHGILSIEQLISRGVGIVGVTLMALLSGFGAVNCPTYMSYFLRNVTDTDLALERRLLQ
TMDMIISKKKRMMAMARRTMFQKGEVHNKPSGFWMGIKSVTTSAGSENLTЛИQQEVDALEELSRO
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGMY
FVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCCTGGTTATGAAGTGGAGCAGTGAGTGAGCCTCAACATAGTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCGCCTCATATCACCAAGGTGGGCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGTGGCCAGGTGCTTCAGGCTTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCTGGTCCAAGGCTTTCGCTGAGAAGAGCTTCCATCCAGGT
GTCATGCGAATTATGGGATCACCCCTGTGAGCAAAAAGGCCAACAGCAGCTGAATTTCACAG
AAGCTAAGGGCCGTAGGCTGCTGGACTAAGTTGGCCAGGACAAAGTGAACACAGCC
TTGAAAGCTAGCTTGAACACTGGCAGCTATGGCTGGGTTGGAGATGGATTGTCATCTCTAG
GATTAGCCCAAACCCAAAGTGTGGAAAAATGGGGTGGGTCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCACTCCAGAA
ATTATCACCAACAAAGATCCCATTCAACACTCAAACACTGCAACACAAACAGAATTATG
CAGTGACAGTACCTACTCGGTGGCATCCCTTAECTCTACAATACCTGCCCCACTACTCTC
CTGCTCCAGCTCCACTTCTAFTCCACGGAGAAAAAAATTGATTGTCACAGAAGTTTATG
GAAACTAGCACCAGTCTACAGAACTGAACCATTTGTGAAATAAAGCAGCATTCAGAAGTGA
AGCTGCTGGTTTGGAGGTGCCCCACGGCTCTGCTAGTGCTGCTCCCTCTTTGGTCTG
CAGCTGGCTTGGATTTCGATTTGCTATGTCAAAAGGTATGTAAGGCCCTTCCCTTAAACAAAGAAT
CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCATGATAGCAACCC
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACACTACCGTGC
GATGCTGGAAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACTGAGGCTGGTTCTT
CATGCTCCTTACCCCTGCCCCAGTGGGAAATCAAAGGCCAAAGAACCAAAGAAGAAAGTCCA
CCCTGGTCTTAACTGGAATCAGCTCAGGACTGCCATGGACTATGGAGTGACCAAAAGGAAT
GCCCTTCCTTATTGTAACCTGCTGGATCCATTCCCTACCTCCAAAGCTTCCACGGCCT
TTCTAGGCTGGCTATGCTTAATAATATCCACTGGGAGAAAGGAGTTTGCAAAAGTGCAGGAC
CTAAAACATCTCATGAGTACAGTGGAAAAAGGCCCTGGCTGCTGAGGCTAGGGTTG
AAAGGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGGAAACACGCTATCCCACCTGACATGCTTCTGGCAGCCGGTAAGGCAAAAAGAAT
GGCAGAAAAGTTAGCCCCCTGAAAGCCATGGAGATTCTCATAACTTGGAGACCTAATCTGTAAA
GCTAAAATAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTCAGCAGGACTGTAAAC
ACAGACAGGGTCAAAGTGTTTCTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGCTCTACCACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAACCTTATAAATTCTATTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGCAAGGTATTACACTGTAATTGAATATTCTCAAAAATTCACATAGTAG
AACGCTATCTGGGAAGCTATTCTTCTAGGTTGATATTCTAGCTTATCTACTTCCAAACTAAT
TTTATTTCTGAGACTAATCTTATTCTATTGCTAATATGGCAACCCATTATAACCTTAATT
TATTATAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAAGCACATTAAAAGTGC
ATTAACAAATGTTACACTAGCCCTCTTTCCAAACAAGAAGGGACTGAGAGATGCAGAAATATT
TGTGACAAAAAAATTAAAGCATTAGAAAACCTT

APP_ID=10063518

FIGURE 6

MARCFSLVLLLTSIWTTTRLLVQGSLRAEELSIVSCKRIMGITLVSKKANQQLNFTAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNNPKCGKNGVGVLIWKPVSQFAAYCYN
SSDTWTNSCIPEIITTQDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFTPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCCGCGCTCCCGCACCCGGGCCCGGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCCCGG
 GGCTCTCCCGGGGGAGCGAGCAGATCCAGTCCGGCCGCAAGCGAACACTCGGTCACTGGGGCG
 CGGCTGCGGGCGCAGAGCGAGATGCAAGCGGCTTGGGCAACCTGCTGTGCTGCTGGCG
 CGGGGGTCCCCACGGCCCCCGCGCCGCTCGACGGCACTCGGCTTCAAGGCCGGCG
 GCTCTCAGCTACCCGCAAGGAGGACCCCTCAATGAGATGTCAGCGGCTTGGGCAACCTGCTGTGCT
 GGAGGACACGCAGCACAAATTGCGCAGCGGGTGAAGAGATGGAGGCAAGAAGCTGCTGCTA
 AACATCATCAGAAGTGAACCTGGCAAACCTACCTCCCAGCTATACAATGAGACCAACACAGAC
 ACGAAGGTTGAAATAATACCATCATGCAACCGAGAAATTACAAGATAACCAACACAGAC
 TGGACAAATGGTCTTTCAAGAGACAGTTATCACATCTGAGGAGACGAAGAAGGCAGAAGGAGC
 ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTAUTGCCAGTTGCCAGTCCAGTAC
 ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGGAGACCA
 GCTGTGTCTGGGTCACTGCACCAAAATGCCACCAAGGGCAGCAATGGACCATCTGTGACA
 ACCAGAGGGACTGCCAGCCGGCTGTGCTGTGCCCTCCAGAGGAGGCCGTGTTCCCTGTG
 ACACCCCTGCCCGTGGAGGGCAGCTTGCATGACCCCGCCAGCCGGCTTCTGACCTCATCAC
 CTGGGAGCTAGAGCCTGATGGAGCCCTGGAGGATGCCCTTGCCAGTGGCTCTGCCAGC
 CCCACAGCCACAGCCTGGTATGTGCAAGCCGACCTCGTGGGGAGCCGTGACCAAGATGG
 GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAGTTGCCAGCTCATGGAGGAGGTGCG
 CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCGTG
 CCGCCGCTGCACTGCTGGAGGGAGAGATTTAGATCTGGACCAGGCTGTGGTAGATGTGCAA
 TAGAAATAGCTAATTATTCCCCAGGTGTGCTTGGCTGACAGCATGGGTGTGCAATTGTCAGCT
 CATCTTCTCCAGTAAGTTCCCTCTGGCTGACAGCATGGGTGTGCAATTGTCAGCT
 CCCCCAGGCTGTTCTCCAGGCTTCAGCTGCTGGCTGGAGAGTCAGGAGGGTAAACTGCA
 GGAGCAGTTGCCACCCCTGTCAGATTGGCTGCTTGGCTCTACCAAGTTGCCAGACAGCCG
 TTGTTCTACATGGCTTGATAATTGTTGGAGGGAGGAGATGAAACAAATGGAGTCTCCCTC
 TGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCAAAACATCACTGGAAAAATG
 CAACAAATGAATTTCACCGCAGTTCCATGGCATAGGTAAGCTGTGCCCTCAGCTGTTG
 AGATGAAATGTTCTGTCACCCCTGCATTACATGTGTTATTCACTCCAGCAGTGTGCTCAGCTC
 TACCTCTGCCCCAGGGCAGCATTTCAATCCAAGATCAAATCCCTCTCTCAGCACAGCCTGGGG
 AGGGGGTCTTGGCTCCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGC
 CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGGTGACTCTAAGCTCAGTGTCT
 CTCCACTACCCACACCAGGCTTGGTGCACCCAAAGTGTCTCCCCAAAAGGAAGGAGAATGGGAT
 TTTCTTGAGGCATGCACATCTGGAAATTAGGTCAAACTAATTCTCACATCCCTCTAAAGTAA
 CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCTTAATGAAGACAATGAT
 ATTGACACTGTCCTTGGCAGTTGCATTAGTAACTTTGAAAGGTATATGACTGAGGCTAGCA
 TACAGGTTAACCTGCAGAAACAGTACTAGGTAATTGTAAGGGCAGGATTATAATGAAATTG
 AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
 TGTGAAACATGGTTGAAATATGCGACTGCGAACACTGAACTCACGCCACTCCACAAATGATG
 TTTCAAGGTGTCATGGACTGTTGCCACCATGTTACATCCAGAGTTCTAAAGTTAAAGTTGCA
 CATGATTGATAAGCATGCTTCTTGAGTTAAATTATGATAAAACATAAGTTGCAATTAGAA
 ATCAAGCATAAAACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPAISYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLNLPPSYHNETNTDKVGNNTIHVHREIHKITNNQTGQMVFSE
TVITSGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQMLCTRSECCGDQLCVWGHC
TKMATRGSGNTICDNQRDCQPGLCQAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACGCGTGGGGCTGTGAGAAAGTCCAATAAATACATCATGCAACCCCAC
GGCCACCTTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCAAAG
GCCTAACCTAACGTTCTGCTTCATCTGAAATCTATGGGGTCTGGGGCTCTCTGGACCCCTT
AACTGGGTACTGCCCTGGCCAATGCGCCTCGCTGGAGCCTTGCCTCCTACTGGGCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTGCCTCATCCGCACACTCCGTACC
ACACTGGGTATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCGGGTACATTTG
GAGTATATTGACCAAGCTCAGAGGAGTGCAGAACCCCTGTAGCCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCCTGGTGTGGAAAAATTATCAAGTTCTAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTGTCTAGCCAAAATGCGTTATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCGTCCCTGGACAAAGTCACAGACCTGCTGCTGTTCTTGGGAAGCTGCTGGT
GGTCGGAGGGCTGGGGCTCTGTCCTTTCTCCGGTCCATCCGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGCTAT
GTCATGCCAGCGGCTTCTCAGCGTTTCGGCATGTGTGGACACGCTTCCCTGCTTCC
GGAAGACCTGGAGCGAACACGGCTCCCTGGACCGCCCTACTACATGCCAAGAGCCTCTAA
AGATTCTGGGAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCTTACAGGT
CTCCATTTGTGGTAAAAAAAGGTTTAGGCCAGGCCGTGGCTCACGCCGTGAATCCAACACT
TTGAGAGGCTGAGGCCGGGATCACCTGAGTCAGGAGTTGAGACCAGCCTGGCAACATGGT
AAACCTCCGTCTCTATTAAAAATACAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCC
GCTACTCGGAGGCTGAGGCAGGAGAATCGCTGAACCCGGGAGGCAGAGGTTGAGTGGCA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAA
AAGATTTTATTAAGATATTTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCKFLRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLLLFFGKLLVVGGVGVLSSFFFFSGRIPGLGKDFKSPHLNYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCCAATGGGGGCTGCCTGGGAGCCTGC
 TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
 CAGCGCAACTCCACCCTGAGCCGCCATCTTCACGTTCCCTTCCCTGGGGTGCTGGTGTCCA
 TCATTATGCTGAGCCGGGGCTGGAGACTCAGCTCTACAAGCTGCCCTGGGTGTGAGGAGGGGCC
 GGGATCCCCACCGTCTGCAGGGCACATCGACTGTGGCTCCCTGCTGGCTACCGCGCTGTCTACCG
 CATGTGCTTCGCCACGGCGCCCTTCTTCTTCAACCTGCTCATGCTCTGCGTGAGCAGCA
 GCCGGGACCCCGGGCTGCCATCCAGAATGGTTTGGTTCTTAAGTCCCTGATCCTGGTGGCCTC
 ACCGTGGGTGCCCTACATCCCTGACGGCTCCCTCACCAACATCTGGTCTACTTCGGCGTCGTGG
 CTCCCTCCTTCTACCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTCTGGAACCAGCGGT
 GGCTGGCAAGGCCAGGAGTGGCATTCCGTGCTGGTACGCCAGGCTCTTCTTCACTCTCCTC
 TTCTACTTGCTGATCGCGGCCGTGGCCTGATGTCATGTAACACTGAGCCCAGCGGTGCCA
 CGAGGGCAAGGTCTCATGCCCAACCTCACCTCTGTGTCTGCGTGTCCATGCTGTGCCCTGC
 CCAAGGTCCAGGACGCCAACCTGGGTCTGCTGCAGGCCCTGGTCATCACCTCTACACCATG
 TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
 GGGCAACGAGACAGTTGTGGCAGGCCAGGGCTATGAGACCCAGTGGTGGATGCCAGCATTG
 TGGCCTCATCATCTTCTCTGTGCACCTCTCATCAGTCTGCGCTCCAGACACCAGGCAAGGTG
 AACAGCCTGATGCGACGCCAGGAGTGCCACCTATGCTAGACGCCACACGCCAGCAGCAGCAGG
 GGCAGCCTGTGAGGCCGGGCTTGACAACGAGCAGGACGCCGTACCTACAGCTACTCTTCC
 ACTTCTGCCTGGTGTGCCCTCACTGCACGTATGATGACGCTACCAACTGGTACAAGCCGGTGAG
 ACCCGGAAGATGATCAGCACGTGACGCCGTGTGGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCT
 CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCGCAACCGCGACTTCAGCTGAGGCAGCCTA
 CAGCCTGCCATCTGGTGCCTCTGCCACCTGGTGCCTCTGGCTCGGTGACAGCCAACCTGCCCTC
 CCCACACCAATGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
 CTTCTAGTCGTAGTGCCTCAGGGTCCGAGGAGCATCAGGCTCTGCGAGGCCCCATCCCCCGCCAC
 ACCCACACGGTGGAGCTGCCCTTCCCTCCCTGTTGCCACACTCAGCATCTGGATGAA
 AGGGCTCCCTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACTCCCACCAAG
 TGGGGCATCCGGCACTGAAGCCCTGGTGTCCCTGGTCACGTCCCCCAGGGGACCTGCCCTCCTG
 GACTTCGTGCCTTAAGACTCTAAGACTTTCTAATAAACAAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCATAAFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQVLVIDFAHAWNQRWLGKAE
ECDSRRAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPS
VGLIIFLLCFLFISLRSSDHRQVNLSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLLWTLVAPLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCTGGGCGGCCAGGAACCACCGTTAAGGTGTCCTCTTAAAGGATGGTGA
 GGTTGGAAAAAGACTCCTGTAACCTCCAGGATGAACCACCTGCCAGAACATGGAGAACG
 CTCTCACCGGAGCCAGAGCTCCATGCTCTCGCAATATCCATTCAACCCACACAA
 CTCATGCCAGGATTGAGTCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
 TTTCTGTTGTTGTCACCTTGACCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
 TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTATGCACTGACTACTATTCTCATAT
 TTTGATATATTCTCTGGCAGTTTCGATTAAAGTGTAAACTTGCATATGCTGTGCAG
 ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAAG
 TGATCCTTCGAAGCTTTCTCAAGGGCTTTGGCTATGCTGCCATATTCATTCA
 CTTGCCTGGATTGAGACGTGGTCTGGATTTCAAAGTGTACCTCAAGCAGAAGAAGAAAA
 CAGACTCCTGATAGTTAGGATGCTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
 GTCAGTTTATTCCCCTCTGAATCCGAAGCAGGACTGAAGAAGCTGAAGAAAAACAGGACAGT
 GAGAAACCACTTTAGAACTATGAGTACTACTTTGTTAAATGTGAAAACCCCTCACAGAAAGTC
 ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTGAATGGTACGTC
 CACTGCTGGTTATTGAACAGCTAATAAGATTATTGTAATACCTCACAAACGTTGTAC
 CATATCCATGCACATTAGTGCCTGCTGTGGCTGGTAAGGTAATGTCATGATTCACTCTCT
 TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCCAATC
 AAAAGACTTAATATATTGAAGTAACACTTTTAGTAAGCAAGATAACCTTTATTCAATTCA
 AGAATGGAATTTTTGTTCATGTCAGATTATTTGTTAACTTAAACACTCTACATT
 TCCCTGTTTTAACTATGCACATGTGCTTTGTCAGATTAAAGTGTAAATAAAATCTG
 ACATGTCAATGTGGCTAGTTTATTGTTCTGATTGTTGCATTATGTGTTGGCTGAAGTGTGGA
 CTTGCAAAAGGGAGAAGAAAGGAATTGCGAATACATGAAAATGTCACCAGACATTGTATTATT
 TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAAATGC
 ACAAAATGACTAAACCATTCAATCATGTTCTTGCAGCCAATTCAATTAAAATGAA
 CTAATTAAAAA

MNHL PED MEN AL TGS QSS HAS LR NI HS IN PTQ L MARI EY EG RE KK G IS DV R RT FCL F VTF D L L F

FIGURE 14

MNHL PED MEN AL TGS QSS HAS LR NI HS IN PTQ L MARI EY EG RE KK G IS DV R RT FCL F VTF D L L F
V TLL WII EL NV N G G I E N T L E K E V M Q Y D Y Y S S Y F D I F I L L A V F R F K V L I L A Y A V C R L R H W W A I A L T T
A V T S A F L L A K V I L S K L F S Q G A F G Y V L P I I S F I L A W I E T W F L D F K V L P Q E A E E E N R L L I V Q D A S E R
A A L I P G G L S D G Q F Y S P P E A G S E E A E E K Q D S E K P L L E L

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTGGGACCCAGGACCCCTCGGGCCGACCCGCCAGGAAGACTGAGG
 CCGCGGCCTGCCCCGCCCCGCTCCCTCGCCGCCGCCCTCCGGGACAGAAGATGTGCTCCAG
 GGTCCCTCTGCTGCTGCCGCTGCTCTGCTACTGGCCCTGGGGCTGGGTGCAGGGCTGCCAT
 CGGGCTGCCAGTGCAGCCAGCCACAGACAGTCCTCTGCACTGCCGCCAGGGGACCACGGTCCC
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTTGAGAACGGCATCACCAGTCGACGC
 AGGCAGCTTGCCGGCTGCCGGCTGCAGCTCTGGACCTGTACAGAACAGATGCCAGCC
 TGCCCAGCGGGCTTCCAGCCACTGCCAACCTCAGCACACTGGACCTGACGGCAACAGGCTG
 CATGAAATACCAATGAGAACCTTCCGGCTGCCAGCTGCCCTGGGAAAGAA
 CGCGATCCGGCACATCCAGCTGGTCCCTCGACACGCTGCCGCCCTCTGGAGCTAACGCTG
 AGGACAACGAGCTGCCAGCTGCCAACCTGGAGCCCTGCGCCTGCGCTGGACCTCAGC
 CACAACAGCTCTGGCCCTGGAGCCCGCATCTGGACACTGCCAACGTGGAGGCGCTGCCG
 GGCTGGCTGGGGCTGCAGCAGCTGGACGAGGGCTTCACTGCCGCTTGCGCAACCTCCACGACC
 TGGATGTGTGCCACAACAGCTGGAGCGAGTGCCACCTGTATCCGAGGCCCTCGGGGCTGACG
 CGCCCTGCCGGCTGGCCGCAACACCCGCAATTGCCAGCTGCCGCCAGGACCTGGCCGGCTGGC
 TGCCCTGCCAGGAGCTGGATGTGAGCAACCTAACGGCTGCCCTGGGAGACCTCTCGGGCC
 TCTTCCCCCGCTGCCGCTGCTGGCAGCTGCCGCAACCCCTTCAACTGCCGCTGGGAGACCTGAGC
 TGGTTTGGCCCCCTGGGTGCCGAGAGCCACGTCACACTGCCAGCCCTGGAGAGACGCCGCTGCCA
 CTTCGCCAACAGCTGGCCGGCTGCTCTGGAGCTTCACTGCCGACTTTGGCTGCCAG
 CCACCAACACACAGCCACAGTGCCACCCAGGAGGGCGTGGTGGGGAGGCCACAGCCTGTCT
 TCTAGCTTGGCTCCTACCTGGCTAGGCCACAGGCCGGCACTGAGGCCCCAGGCCACCTGCCCT
 CACTGCCAACCGACTGTAGGGCTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCCTCA
 ATGGGGGACATGCCACCTGGGACACGGCACCCACTGGCTGCTTGTGCCCGAAGGCTTCAG
 GGCTGTACTGTGAGAGCCAGATGGGCAGGGGACAGGCCAGCCCTACACCACTGACGCCGAG
 GCCACCACGGCTCTGACCCCTGGCATCGAGCCGGTGAAGCCCCACTCCCTGCCGTGGGCTGC
 AGCGCTACCTCCAGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCACTATCGCAACCTATCG
 GGGCTGATAACGGCTGGTGAACGCTGCCACTGCCGCTGGCTGAGTACACGGTACCCA
 GCTGCCGCCAACGCCACTTACTCCGTGTGTCATGCCCTTGGGGCCGGGGTGCCTGGGAGG
 GCGAGGAGGCCATACACCCCAAGCCGTCCACTCCAACCACGCCAGTCACC
 CAGGCCGCCAGGGCAACCTGCCGCTCTCATGGCCGCCCTGGCCGGTGTCTCTGCCGC
 GCTGGCTGCCGGGGCAGCCTACTGTGTGCCGGGGGGGGGGCCATGCCAGGCCAGGGCTCAGG
 ACAAAAGGGCAGGTGGGGCAGGGCTGGGGCCCTGGAACCTGGAGGGGAGTGAAGGTCCCCCTGGAG
 CCAGGCCGAAGGCAACAGAGGGCGGGTGGAGAGGCCCTGCCAGGGGTCTGAGTGTGAGGTGCC
 ACTCATGGGCTTCCCAAGGCCCTGGCCTCAGTCACCCCTCACCAAGGCCCTACATCTAAAGCCA
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTAGCCAGTGAAGATGCCAGGCCCTCCCTGCTGCC
 ACACCACTGTAAGTCTCAGTCCCAACCTCGGGGATGTGTGCAAGACAGGGCTGTGACCACAGCT
 GGGCCCTGTTCCCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGCCAGCTGACGAGCC
 CTAACGTCCCCAGAACCGAGTGCCTATGAGGACAGTGTGCCCTGCCCTCCGCAACGTGAGTC
 CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACCGCATGCCCTGGGTCTGCTGGGCTCTCCAC
 TCCAGGCCGGACCCCTGGGGGGCAGTGAAGGAAGCTCCCAGGAAAGAGCAGAGGGAGAGCGGGTAGGC
 GGCTGTGACTCTAGTCTGGCCCCAGGAAGCGAAGGAACAAAAGAAACTGAAAGGAAGATGC
 TTAGGAACATGTTGCTTTTAAATATATATTTAAGAGATCCTTCCCATTATTCT
 GGGAGATTTCAAACCTCAGAGACAAGGACTTGGTTTGTAAAGACAAACGATGATATGAA
 GGCTTTGTAAGAAAAAATAAAGATGAAGTGTGAA

FIGURE 16

MCSRVPLLLPPLLALLGPGVGCPGCGCQCSQPQTFTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTDLRLLKLQDNELRALPPLRLPRLLLDDSHNSLLALEPGIILDTANVE
ALRLLAGLGLQQQLDEGLFSRLRNLDHLDVSDNQLERVPPVIRGLRGLTRRLLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLAPGDLSGLFPRLRLLAAARNPNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLLEDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPVTGPVPQPQDCPPSTCLNGGTCHLGRHHLACLCPEGFTGLYCESQMGGQTRPSPTP
VTPRPPRSLTGLIEPVSPSTSRLVGLQRYLOGSSVQLRSLRITYRNLSGPDKRLVTLRLPASLA
TWTQLRPNATYSVCVMPPLGPGRVPEGEAACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAA
LLAALAAGAAYCVRRGRAMAAAQDKGQVPGPGAGPLEGKVPLEGPKATEGGGEALPSGE
CEVPLMGFPGPGLQSPHLAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131 337-345

N-myristoylation sites

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCAGGCGGCGGTGGCTGAGTCCTGGTGGCAGAGCGAAGGCACAGCTC**ATGCG**
GGTCCGGATAGGGCTGACGCTGCTGCTGTGCGGGTGCCTGAGCTGGCCTGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTGATTCAAAGAAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACAGCTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGAAAGATGCCAGACTGTGGTGTG
CTACAACTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTACACTGGAATGAAAATCTTAATGGAAG
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTATTGGTGAATTACTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTGAGACTGAGGAGGCTCTCCAAGGGACAGACTGCTCTGG
CTTTCTGTATGCCCTGGACTGGTGTAAATTCAAGTCAGGCAAAGGCTTGTATATTACAT
TTGGAGCTTGGGGCAATCTAACAGCCCACATGGTTGGTAAGTAGACTT**TAGTGGAAAGGCT**
AATAATATTAACATCAGAAGAATTGTTGTTATAGCGGCCACAACCTTTTCAGCTTCTGATC
CAGATTGCTTGTATTAAGACCAAAATTCAAGTGAACCTCCCTCAAATTCTGTTAATGGATAT
AACACATGGAATCTACATGTAATGAAAGTTGGTGGAGTCCACAATTTCCTTAAATGATTAG
TTTGGCTGATTGCCCTAAAAAGAGAGATCTGATAAAATGGCTTTAAATTCTGAGTTG
GAATTGTCAGAATCATTTTACATTAGATTCTAAATTAAAAATTCTTCTTAGTTTCA
AAATTGGTAAATGGTGGCTATAGAAAAACACATGAAATATTACAAATTGGCAACAATGC
CCTAAGAATTGTTAAATTCACTGGAGTTGGTGCAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTTCTGATGGCTGTCTCCCATTTATTCTGGTCAATTGTTAGTGCAGACTGT
GCCCTGCTTCCAGTAGTCTCATTTCCCTATTGCTAATTGTTAGTCTTGTCAATTGG
AAGATTAACTCATTTAATAAAATTATGCTAAGATTAAAAAAA
AA

MRVRIGLTLCAVLLSLASASSDEEGSQDES LDSKTTLSDESVKDHTAGR VVAGQI FLDSEES EL

ESSIQEEEDSLKSQE GE SVTEDIS FLES PN PEN KDYEPKKV RKP ALTAIE GTAH GEPCHFPFLFLDK
EYDECTS DGDREDGRLW CATTYDYKADEKWGFCE TEEEAAKRRQM QEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERV SYALLFGDYL P QNIQA REMF EKLTEEGSPKG QTALGFLYASGLGVN
SSQAKALVYYT F GALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCA GATTTAAGCCC ATCTGCAGTGGAA TTTCATGAA CTAGCAAGAGGACACC ATCTTCTT
GTATTATA CAAGAAAGGAGTGTA CCTATCACACACAGGGGGAAAAATGC TTTGGGTGCTAGG
CCTCCAATCCTCTGTGGTTCTGTGGACTCGTAAGGAAA ACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGTTGGAAACTTGGCAGCCAGAAC TTTGAT
AAAAAGGGATTCATGTAATCGCTGCCGTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTCTGGATGTGACCGACCAGAGAAC TGTCAAGAGGACTG
CCCAGTGGGTGAGAACCAAGTTGGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCAGCGTGTGGCTCCC ACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTTCCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATAATGTCTCCAGTGTGGAGGTGCCCTGCAATCGTGGAGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAACAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAACAAATGGAGAACGGTTACATTGAAAAA
AGTCTAGACAAACACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAAGTCTCTCCCTAACAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCAAAGGAGTGTGACTCAGCTAACCAACAAATGTCTCCAGGCTATGA
AATTGGCCGATTTCAGAACACATCTCCTTCAACCCCATCCTTATCTGCTCCAACCTGGACT
CATTAGATCGTCTTATTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGTATCCAGGGT
CCCTGCTCAAGTTCTTGAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCCTGCCCT
GTATTTAGGTTGCCTGCTTGGTGTGATGTAAGGAAATTGAAAGACTTGGCCATTCAAATGA
TCTTACCGTGGCCTGCCCATGCTTATGGCCCCAGCATTACAGTAACCTGTGAATGTTAAGT
ATCATCTTTATCTAAATATTAAGATAAGTCAACCCAAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLLEDY
REPIEVNLFGlisVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEFYIEKSLDKLGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGCTGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGCTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTCTCTGGGAAGTAAAAA
GGTGAAGCCAAGAACAGCATTACTGATCCCAAATGGATGATGTTGAAGTTGTTATAACATTGA
CATTAGAAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTTGGTAGGTTGGTACAAATTCCGT
CGTCATTAGCATGACATGAGCTTACAGGAGAGGCTGCTCACAAAACCTGCAGGAGCATT
TTCAAACCAAGACCTGTTCTGCTATTAAACACCAAGTATAATAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCTTATATAACCTCAAAAGGACTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGGATGTCATGAAACAACGTTATAAAACTGTATCAGGTTCTGTATGTCACACTGG
TTTAGCCGAGCAGTACAACACACAGCTCAAATTTTGAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAAACTAGTAAAGGATGTAACACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGGAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA
ACATTTTCTTGTCAAGGATTACGGACCTTTTCAAATTCTGAATTCTTCATTGATGTGTT
ATGCTTTAAAAAAATAGACATGTTCTAAAGTAGCTGTAACTACAACCACATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATCTGGTAGTGTAACTACAAGGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAGATGAAGGGTTGGTGAATTACGGTCTC
CTACATTTGATCTTTAACCTTACAAGGAGATTTTTATTGGCTGATGGTAAAGCCAAAC
ATTCTATTGTTTACTATGTTGAGCTACTGCAAGTTGAGCTAGTACACAG
TGTTGAGCTAATACACAGATAACTCTAGTCATTACTTCACAAAGTACTTTCAAACATCA
GATGCTTTATTCCAAACCTTTTCACTAAGTTGAGCTAGGAGTTAGAGACCAAGCTGGCAACGTATT
ACACATTCTTAAAGGAAAGTGAAGGAGCAGGACAGTGGCTCACACCTGAATCCACAG
TAGGAAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAAGCTGGCAACGTATT
GAGACCATGTCATTAAAAAAATGGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATGAAATTTCTGAGTCATTAAACCTCTTAAAGTGTAACTTTTAAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAAAAAAAAA

PCYQLFSFYNSSGEVNEQALKKILSNVKNNVGWYKFRRHSDQIMTFRERLLHKNLQEHSNQDL
VFLLLTPSIIITESCSTHRLEHSLYKPQKGLFHRVPVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFeedGSLKEVHKINEEMYASLQEELKSICKVEDSEQAVDKLVKDVRNLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNHLDVVDNLTL
MVEHTDIPPEASPASTPQIIKHKALDDLDRWQFKRSRLLTDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCCGGCGGAGGGCAGAGTCAGCCAGGCCAGTCCAGGCCAGGCCAGGCCAGGCCAGGCCAA
GCAGCGCGCAGCGAACGCCGCCGCCACACCCCTCTGCGGTCCCCGCCCTGCCACCCCTCCCTCCCTCCC
GCGTCCCCGCCCTGCCGGCAGTCAGCTTGCCGGTTGCCTGCCCGCGAACACCCCGAGGTACCCAGGCCGCCCTCT
GCTTCCCTGGGCCGCCGCCCTCCACGCCCTCTCCCTGCCCGGCCCTGGCACCGGGACC GTGCCCTGA
CGCGAGGCCAGCTACTTTGCCCGCCTCCCTGCCCTGCCCTTCCACCAACTCCAACCTCCCTCCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTGCCCGCTGCCGTAGCGCGCTCCCGTCCGGTCCAAA
GGTGGGAAACGCGTCGCCCGCCACCAATGGCACGGTTCGGCTGCCCGCTCTGCACCCGGCAGTGCTC
AGGCCCGCCTGGCTGCCAGCTCAAGCTCAAAGTTGCTCGGAAGTGCGACGTCTTACGTGTCAAAGGCTTC
AACAGAACGATGCCCTCCACGGAGATCAACGGTGAATTTGAAGATCTGCCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACGCCGCAAAGTAAGATGATTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTGATGAATTCTCAAAGAAACTACTTGAAAATGCCAGAGAAATCCCTG
AATGATATGTTGTAAGACATATGGCATTATACATGCCAAATTCTGAGCTATTTAAAGATCTCTCGTAGAGITG
AAACGTTACTACGTGGTGGAAATGTGAACCTCGAAGAAATGCTAAATGACTCTGGCTGCCCTGGAGCGGATG
TTCCGCCGGTGAACTCCAGTACCACTTACAGATGAGTATCTGGAATGTTGAGCAGTATACGGAGCAGCTGAAG
CCCTCGGAGATGCCCTCGCAAATTGAAGCTCCAGGTACTCGTGTGTTGAGCAGCCGTACTTCGCTCAAGGC
TTAGCGGTTGCCGGAGATGTCGTGAGCAAGGTCTCGTGTAAACCCACAGGCCAGTGTACCCATGCCCTGGTAG
ATGATCTACTGCTCCACTGCCGGGCTCGTACTGTGAAGGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTTGGCAACCAAGGGATCTGATTGAAATGGAACAATTCTAGATGCTATGCTGATGGTGGCAGAGGGCTA
GAGGGTCTTCACATTGAATCGGTATGGATCCCCTCGATGTAAGATTCTGATGCTATTGAAACATGAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTTGACCCCTCCAGCTGGACGAATTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGTGCTCAGACACATCAGGCCAGGAAACGCCAACAGCAGCTGGCACT
AGTTGGACCGACTGGTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAAATTCTGGTCTCCCTCCAGCAAC
GTTTGCACGATGAGAGGATGGCTGCAGGAAACGCCAATGAGGATGACTGTTGGAATGGAAAGGCAAAGCAGGTAC
CTGTTTGAGTGCAGGAAATGGATTAGCAACAGGGCAACACCCAGGGTCAGGTGACACCAGCAACCCAGAC
ATACTGATCCTCGCAAATCATGGCTCTCGAGTGTGATGACCAGCAAGATGAAAGATGCAATACAATGGAAACGACGTG
GACTTCTTGATATCAGTGTGAAAGTAGTGAGAAGGAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGAATGCAACTGACATGACCATGCTGGAAAGGTGCAATGAGAAAGGCCACAGTGTGGTGTCCGTCTGGGCA
CAGGCCCTACCTCCTCACTGCTCTGCTCATCTGGTTATGCGAGAGAGTGGAGATAATTCTCAAACCTGAG
AAAAAGTGTCAATGAAAGTAAAGGCAACAGTATCACTTTCTACCATCTAGTGTGACATTGTTTAAATGAA
TGGACAACATGTACAGTTTACTATGTTGCACTGGTTAAGAAGTGTGACTTTGTTCTCATTGAGTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGTTCTGCTCCCCAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTGTGATTATCACTCTATTGTTGTTGATGTTTCTCATTGTTGTTGGGTT
TTTTTTCAACTGTGATCTGCCCTGTTCTAACAGCAAACCAAGGGTCCCTTGGCACGTAACATGTACGTATT
TCTGAAATTAAGTGTACAGAAGCAGGTTTATTTATCATGTTATCTTAAAGAAAAGCCAAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKS KSCSEVRRLYVSKGF N KNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSI L QSKDDFKSVVSEQC NH LQAVFASRYKKFDEFFKELLENAEKS LNDMFVKTYGH
LYMQNSELFKDLFVELKRYYYVGVNVNLEEMLNDFWARLLERMFR LVNSQYHTDEYLECVSKY TE
QLKPF G D V P R K L K L Q V T R A F V A A R T F A Q Q L A V A G D V V S K V S V V N P T A Q C T H A L L K M I Y C S H C R G L
VTVKPCNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDP1DVKISDAIMN
MQD NS V Q V S Q K V F Q G C G P P K P L P A G R I S R S I S E S A F S A R F R P H P E E R P T A A G T S L D R L V T D V K
EKLKQAKKF W S S L P S N V C N D E R M A A G N G N E D D C W N G K G K S R Y L F A V T G N G L A N Q G N N P E V Q V D T S
KPDILILRQIMALRVMTSKMKNAYNGNDV D F F D I S D E S S G E G S G S G C E Y Q Q C P S E F D Y N A T D H A G
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glycans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCCAGGAGGCCACAGGGACCAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGCGGCCAAGAATGTGAGTGCAAAGATTGGTCTGAG
AGCCCCGAGAAGAAAATTCTAGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACACAGAAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTTGTAGGAGCTCTG
AGC GCCC ACTCTCCA ATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTAACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAAATATTTT
AAATGTCAAAAAAAAAAAAAAA

MKVLISLLLPLMLMSMVSSSLNPVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM

FIGURE 26

MKVLISLLLPLMLMSMVSSSLNPVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPKCDHKGNVKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
 AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCTGCAGCTGCTGGTGTGCTTACCTGCC
 CCTGCACCTCATGGCTCTGCTGGCTGTCAGCCCTGTGCAAAAGCTACTTCCCTACCTGA
 TGGCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAACGGGAGCTTCAGCCAG
 ATAAAGGGCTTACAGGAGCCTCCGGAAAGTGGCCTACTGGAGCTGGCTGCGGAACCGGAGC
 CAACTTCAAGTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTGAGA
 AGTTCCGTACAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGCTCCT
 GGAGAGGACATGAGACAGCTGGCTGATGGCTCATGGATGTGGTGTGCACACTGTGGCTGTG
 CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
 TCTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGCCTCATGTGGCAGCAAGTTTC
 GAGCCCACCTGAAACACATTGGGGATGGCTGCTGCCCTCACAGAGAGACCTGGAAGGATCTGA
 GAACGCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
 GGCCCCACATCATGGGAAAGGCTGTCAAACAAATTTCCAAGCTCCAAGGCACACTTGTCTCC
 TTCCCCAGCCTCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACTA
GCAGAACATGAGAGAACATTCACTGTACCACTACTAGTCCCTCTCCCCAACCTGTCCAGGG
 AACACTAGGACCCCTGTTGATCCTCAACTGCAAGTTCTGGACTAGTCTCCAACGTTGCCTC
 CCAATGTTGTCCTTCCCTCGTTCCCATGGTAAAGCTCCTCGCTTCCCTGAGGCTACAC
 CCATGCGTCTCTAGGAACGGTACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
 CCTCTCTCCCCACTACCACCTTCCCTGAGCTGGGGCACCAGGGAGAACATCAGAGATGCTGGG
 ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCAAATATTTTAATAAATAGACGAA
 ACCACG

MDILVPLLQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELSQIKGL

TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRLRPGGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATTTGCCTATCCACCTCCCCAAGCCCCTTAACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGTGCCCAGAAAGTCTCTTCTG
CCACTGACCCCCATCAGGGATTGGGCCTTCTTCCCCCTTCCTTCTGTCTCCTGCCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGGGATGGC
AAGAAAGCTGGGAGATAGGGAACAGAACAGAGGGTAGTGGGTGGGCTAGGGGGCTGCCTTATTAAA
GTGGTTGTTATGATTCTTACTAATTATAACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTCCCCGTGTTCAATGTTGTAAGATTGTTCTGTGTAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTCGATCTGTTGATAAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGGAAATCTGTCGG
GAAGAGATAACAATCCTGGCCTGTGTACCTCGCATAGCCTTGCTTGGCCATGATGTTACC
TTCAGATTCTACCAACCCCTCTGGTACATTTCATTGTTATTGGGATTGTTGTT
TGTCTGCGGTGTTATGGTGGCTGTATTGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTGATTTTGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGGCTCTGGTGGCTGTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGCCAAGTGGAAATATAAGCCCTTCCGGCATTGGTACATGTGGTGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCTACCCCTGCTGAGCAGAAATGACTATAGCTGGGAG
TGGTTACTTGTATTCACAGAAGTAAAATGATCCTCCTGATCATCCATCCTTGTCTCTC
TCCATTCTCTTCTTACCATCAAGGAACCGTTGTAAGGGTCAATTAAATCTGTGGTGTAC
GATTCCGAGAACATGTGTCATGCTACATGCAAAACGCACTGAAAGAACAGCAGCAGTGGTGTAC
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTCTGGTGTCTGACAATACCTGCTCCATCTC
AACCAGAACATGCAACTACAACTGCTATTAAATGGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAAATCTGTCCAAGAACACTCAAGTCACCTTACATCTATTAACTGCTTGGAGACTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTACTGTTGGAGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGGCAGTCCTCTGTTATTGGTAGCTTGGCTACTTAGTAGC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAACATTCTGAGTTGATC
AAAAGGAGCAACAAATTAAACATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACCTCCAGGCCATTGTGAGATAGAACCCATTAGGTATCTGTACCTGGAAACATT
TCCTTCTAAGAGCCATTACAGAATAGAAGATGAGACCACTAGAGAAAAAGTTAGTGAATT
TTAAAAGACCTAATAAACCTATTCTCCTCAAA

FIGURE 32

MSGRDTI~~L~~GLCILALALS~~L~~AMMFTFR~~F~~ITLLVHI~~F~~ISLV~~I~~LG~~L~~FVCGVLWWLYYDYTN~~D~~LSIE
LDTERENMKCVLGFAIVSTGI~~T~~AVLLVLI~~F~~VLRKRIKLTVELFQITNKA~~I~~SSAPF~~L~~FQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVM~~E~~GGQVEYKPLSGIRYMW~~S~~YH~~L~~IGLIWTSEF~~I~~LA~~C~~QQMTIA
GAVVTCYFNRSKNDPPDH~~P~~ILSSL~~S~~ILFFYHQGTVVKG~~S~~FLISVV~~R~~IPRII~~V~~MYMQNALKEQQHG
ALSRYLFRC~~C~~CYCCFWCLDKYLLH~~L~~NQNAYTTAINGTDFCTS~~A~~DKILSKN~~S~~HFTSINC~~F~~GD
FII~~I~~FLGKV~~L~~VVCFTVF~~G~~GLMAFNYNRAFQVWAP~~L~~LLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAAAAGGTCTTGGACCTCTCCCTGTTCTCTTAGA
ATAATTGTATGGGATTTGTGATGCAGGAAAGCTAACGGAAAAGAATATTCAATTGTGTTGGT
GAAAATTTTGAAAAAAATTGCCTCTTCAAACAAGGGTGTCAATTCTGATATTATGAGGAC
TGTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTGGTGTACTGGAGTAC
ATTCAAACAAAGAACGGCAAAGAAGATTAAAGGCCAAGTTCACTGTGCCAGATCAACTGC
GATGCAAAGCCGAAAGATCATCGATCCTGAGTTCAATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAAGGAGGGAAAATACTGTGTCGGAAGGTTGCTGGACAGTCT
GGTACAAAGGGAGTATTCCAACGGTGTCAAATCGTTACCCATACAGATGGAGAGAATCCTT
TATCGTCTTAAAGTAAACCCAAAAGGGTGTAAACCTACCCATCACGCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAAGCTATCAGAGGCCACCTATTCCAGGG
ACAATGCAAGCCGGTACTCTGATGCAGCTTCTGGCTGTACTGTAGTGTGGCCACCCCAC
CACCTTGCCAAGGCCATCCCTCTGCTGTTCTACCACCAGCATCCCCAGACCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTATCAAAGGCAAGATCCTCAGGAGCTGCCCTCAGAAACCTGTGGAGC
GGATGTCAGCTGGGACTTGTCCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC
TGGAGATCCAACACTGCAAATTGACTGTGCTTTAAATTGATGGAGCACCAGCATTGGCAA
CGGCATTCGAATCCAGAAGCAGCTCTGGCTGTGTCGTTCTGGCCAGCTTGTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCAGTGGAGACAACTCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATTCTGAGATCTGAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
GTAGGTGGGCCATCTCCTTGTGACCAAGAACTTCTTCCAAAGCCAATGGAAACAGAAGCGG
GGCTCCAATGTGGTGGTGGATGGGGATGGCTGGCCACGGACAAAGTGGAGGAGGCTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTCATCACCATTGAGGTGCTGCTGAAATGAG
AAGAGTATGTGGTGGAGCCAACTTGCAAACAAGGCCGTGTCAGAACAAACGGCTTCTACTC
GCTCCACGTGAGAGCTGGTTGGCCTCCACAAGGCCCTGCAGCCTCTGGTGAAGCGGGCTGC
ACACTGACGCCCTGGCCTGAGCAAGACCTGCTGAACTCGGCTGACATTGGCTCGTACGAC
GGCTCCAGCAGTGGGGACGGGCAACTTCCGACCCGCTCCAGTGTGACCAACCTCACCAA
AGAGTTTGAGATTCCGACACGGACACGCACATGGGGCGTGCAGTACACCTACGAACAGCGC
TGGAGTTGGGTCGACAAGTACAGCAGCAAGCCTGACATCTCAACGCCATCAAGAGGGTGGC
TACTGGAGTGGTGGCACCGACCGGGGCTGCCATCAACTTCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCAGCAGGGAGGCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATGCGATAGGCGTTGCTGGGCT
GCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGGACCACTCCTTCTTGTGGACGA
GTTGACAACCTCCATCAGTATGCCCCAGGATCATCCAGAACATTGTCAGAGTTCAACTCAC
AGCTCGGAACTGAATTCAAGAGCAGGAGCAGCACAGCAAGTGTGCTTTACTAAGTGTGTT
GGACCAACCCACCGCTTAATGGGGCACGGCACGGTGCATCAAGTCTGGGAGGGCATGGAGAAC
AAATGTCTTGTATTATTCTTGCATCATGCTTTCATATTCCAAAATGGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCAAAGGCTACATCATGTTGAGGGTGTGGAGATTTCAT
TTTGACAATTGTTCAAATTAAGTGTGGAATACAGTGCAGCCCTTACGACAGGCTACGTAG
AGCTTTGTGAGATTAAAGTGTATTCTGATTGAACCTGTAAACCTCAGCAAGTTTCAT
TTTGTCTGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAAAATAAAAAAAA
AAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTVPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPPTLPRPSPSASTTSIPRPQSVGHRSQEMDLWSTATYTSSQRPRADPGIQRQDPSGAAFQKP
VGADVSGLGLVKEELSTSLEPVSLGDPNCIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKTHTNCSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLQPLVKRVCDTDLACSKTCIISADIGFVIDGSSSVGTGNFRVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRKGVYWSGGTSTGAAINFALQL
FKKSCKPNKRKLMLITDGRSYDDVRIPAMAHLKGVITYAIGVAWAQQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCCTGCCGTTAGGAGGTGGCTGCCGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTCTGTTTCAGAGTAGTTACAACAGATCTGAGTGTTAATTAAGCATGGAAT
ACAGAAAAACAACAAAAACTTAAGCTTAATTCATCTGGAATTCCACAGTTCTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTACCGTGGTGCTCCGACTACTCACCCCGAGTGTA
AAGAACCTCGGCTCGCGTCTGAGCTGCTGTGGATGCCTCGGCCTCTGGACTGTCCTCCGAGTA
GGATGTCACTGAGATCCCTAAATGGAGCCTCGCTGCTGTCACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACCGCGTAACCGATGTAATTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTCACCTCACACTCGAGAGCATTCAAACGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCACCCTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGTGAAAAAAAGTCT
TGGTGGGATATGAGGTTCTTACATTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAATGTTGC
ATTGTCCTTAGAGGATGAACACCTCTTATGGTACATAATCCGACAAGATTTTAGACACATATAATA
ACCTGACCTTGAAAACCATTATGGCATTAGGTGGGTACTGAGTTGCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATTTCATCAAACTGGCAATTAGTGAAGTATCTTTAAACCTAAACCACTCAGA
GAAGTTTCACAGGTATCCTCTAATTGATAATTATCCTATAGAGGATTTACCAAAAAACCCATATT
CTTACCAAGGAGTATCCTTCAAGGTGTTCCCTCATACTGCAGTGGGTGGTTATATAATGTCCAGAGAT
TTGGTGCAAGGATCTATGAAATGATGGTACCGTAAACCCATCAAGGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAGTGAACATTCAATTCCAGAACACAAATCTTCTATAGAATCC
ATTGGATGTCCTGCAACTGAGACGTGTGATTGCAGCCATGGCTTCTCCAAGGAGATCATCACTTT
TGGCAGGTATGCTAAAGAACACCACATGCCATTATTAACTTCACATTCTACAAAAAGCCTAGAAGGACAG
GATACTTGCGAAAGTGTAAATAAGTAGGTACTGTGGAAATTCAATGGGAGGTTTGTCAAAGAAATTAAAG
ACCAACAAATTGGACATGTCATTCTGTAGACTAGAATTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAACATGTAGAGTTATTGAAACATTGACTGACTGAAAGGTTGTGA
TATCTTATGTGGATTACCAATTAAAGTATGAGTTCTGTGTCAAAAAAACTTCTCACTGAAGTTATA
CTGAACAAATTTCACCTGTTGGTCATTATAAGTACTTCAGATGTTGCAGTATTCACTGAGTTATT
ATTATTAAATTACTTCACCTTGTGTTTAAATGTTGACGATTCAATACAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAACGATGTCAGTTATTGATACATCACTCCA
TTAATGTAAGTCATAGGTCAATTGATCACTGAAATCTCTGGACTTGTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

© 2024 Biotechnology & Applied Biochemistry

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIHQDFLDTYNNNLTKTIMAFRWTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSDLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAAGCCGGCAAGGATGGAGCTGGGTGCTGGACGCAGTTGGGCTCACTTTCTTCAGCTCCTTCATC
TCGTCCTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGACTGCTGTG
AATATGATCAGATTGAGTGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGACTGCTGTG
TGAGTGTGACTCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGACTGCTGTG
ACCTTGGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCAGAGTGGAGCAGAGTGGAAATATCATGTGTCGGGACTGCTGTG
GCCAGGTTCTGCAGGCCCAAAGGGTCAGATTTGGAAAGCTATCCCCTAACATGTGTCAGAGTGGAGCAGAGTGGAAATATCATGTG
TAAACCTGGGTTGTCACTCCAACTAAGATTGTCACTGTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGAGTGGAGACAACCGCGATGCCAGATCATCAAGCTGTCTGTGGCAACCAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCCACTCCACGCCCTTCCACTCGATGGCTCCAAGAATTGACGGTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCTCATCCCCCTGTTCCATGACGCCACGTGCTGCTTGACAAGGCTGGATCTTACAAGTGTGCCCTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCTTGAAGAAAAGAACTGTCAGACCCCTGGGAGCAGTCATGGTACAGAAAA
TAACAGGGGCCCTGGCTTATCAACGGACGCCATGCTAAATTGGCACCGTGGTCTTGTAAACAACCTCATGT
TCTTAGTGGCAATGAGAAAAGAACTGCCAGCAGAATGGAGAGTGGTCAAGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTCAGACCTGGTGAAGAGAGTCTTCCGATGCCAGGTTCACTGCAAGGGAGACACCATTACACAGCTAT
ACTCAGCGCCCTCAGCAAGCAGAAACTGTCAGAGTGCCTTACCAAGAAGCCAGGCCCTTCCCTTGGAGATCTGCCATGG
ATACCAACATCTGCATACCCAGCTTCACTGAGTCATCTCACCTCTACCGCCCTGGCAGCAGCAGGAGACATGT
CTGAGGACTGGGAAGTGGAGTGGGGGGCACCATCTGCATCCCTATCTGGGGAAAATTGAGAACATCAGTGTCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCCATCACAGGAGGACCAGCGGGGTCATGCCAGGCCCTACACAAGGGAGC
GTGGTCCCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGACTGTGGTGGCTGCCACTGTGTTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACTGAAATTGTTGGGAAATTCTACCGGATGATGACCGGGATGAGAACGACATCC
AGAGCCTACAGATTCTGCATCTGCATCCCAACTATGACCCCATCTGCTGATGCTGACATGCCATCTGAAGCT
CCTAGACAAGGCCCTATCAGCACCCGAGTCCAGGCCATCTGCCCTGCGCTGCCAGTGGGATCTCAGCACTCCCTCCAGGAG
TCCCACATCACTGGCTGGCTGGATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTGGGG
TGGTCAGTGTGGTGAACCTGCTGCTGTGAGGAGCAGCATGAGGAGGACATCCAGTGAAGTGTCACTGATAACATGTT
CTGTCAGTGTGGTGAACCTGCTGCTGTGAGGAGCAGCATGAGGAGGACATCCAGTGAAGTGTCACTGATAACATGTT
CGAGCATCTCTGAGCCACGCTGCCATGTGATGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTCACCAAGGTGCTGCCCTTAAAGACTGGATTGAAAGAAATATGAAATTGACCATGCTCATGCCACTCCCTGAGAAG
TGTTCTGTATATCCGCTGTACGTCATTGCCGTGAAGCAGTGTGGGCTGAAGTGTGATTGCGCTGTGAACCTGGCT
GTGCCAGGGCTCTGACTTCAGGGACAAAACAGTGAAGGGTGAAGTAGACCTCATTGCTGGTAGGCTGATGCCCGTCCA
CTACTAGGACAGCCAATTGGAAAGATGCCAGGGCTGCAAGAAGTAAGTTCTTCAAGAACGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGTCTTCCCAACTTCACTGAGTATACGAATGCCATCACCTGGCAGGAGATCTGGCTTCAAGAACG
GCCCTTGTGAGGCTCTCAAGTTCAAGAGCTGCCAGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCT
TGTGACATGGCACAGTACAGTCTGGCTTCTCCCTTCCCATCTTGATCACATTAAATAAGGTTGGCTTCAAGAACG
GAACATACAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISLSPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSGWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRGENLLEERNCSDPGGPVNQYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCCQONGEWSGKQPICIKACREPKISDLVRRRVLPMQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMDGYQHLHTQLQYECISPFYRRLGSSRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAFLVCAGLVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDDEKTIQSLQISAIILHPNYDPILLADIAILKLLDKARISTRVQPICLAASR
DLSTSFAQESHTIVAGWNVLADVRSPGFKNDSLRSQVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSIDCTAETGGIAAVSFPGRASPEPRWHLMGLVWSYDKTCSHRLSTAFTKVLPDFDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCATACCTCTCATCTGAGAATCAGAGAGCATAATCTTACGGGCCGTGATTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTGGCTGG
TTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGA**ATGAAGGC**GCTCTGTG
TGGCTTGCTTGGCTCAGCTGCTAACTACATTGACAATGTGGCAACCTGCACTTCTGTATTAGAACTCTGTA
AAGGTGCCCTCCACTACGGCCTGACCAAAGATAGGAAGAGGCAGCTACAAGATGGCTGTCCAGACGGCTGTGAGGCC
TCACAGCACGGCTCCCTCCCAGAGGTTCTGCAGCTGCCACCATCTCTTAATGACAGACGAGCTGGCTAGACA
ACCCCTGCTACGTGCTCTCGGCAGAGGACGGGAGCCAGCAATCAGCCCAGTGGACTCTGGCGAGCAACCGAAGTA
GGGCACGGCCCTTGAGAGATCACTATTAGAACGAGATCATTAAAAAAATAATCAGGCTTGAGTGTCTCGAA
GGACAAAAGCGGGAGTGCAGTTGCCAACCATGCCGACCAAGGGCAGGGAAATTCTGAAAAACACCACTGCCCTGAAG
TCTTCCAAGGTTGATACCACCTGATTCCAGATGGTGAATTACCAAGCATCAAGATCAATCGAGTAGATCCAGTGAA
GCCTCTCTATTAGGCTGGTGGAGGTAGCAGAACCCCCACTGGCCATATCATTATCCAACACATTATCGTGTGGGG
TGATCGCCAGAGACGGCCGCTACTGCCAGGAGACATCATTCAAAGGTCAACGGGATGGACATCAGCAATGCC
ACAACATCGCTGCGCTCTCTGCCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCCGAAACAGAAGTCCGCA
GCAGGAACAATGGCAGGCCCGGATGCCCTACAGACCCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGGAGCAGTTGAATAAAACTGGTGCAGAACAGTGGTGGATGAGCCTGGGTTTCATCTCAATGTGCTGGATGGCG
GTGTCGATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATTCGATATGGCA
GCCAGAAAAGTGGCTCATCTGATTAGGCCAGTGAAGACGTGTTACCTCGCTGTCCGCCAGGTTCCGAGC
GGAGCCCTGACATTTCAAGAACAGCCGCTGGACAGCAATGGCAGCTGGTCCCCAGGGCAGGGAGAGGAGCAACA
CTCCAAGCCCCATCCTACAATTACTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCGGTGAATCTCTCG
GCATGACCGTCGAGGGAGCATCACATAGAGAATGGATTGCTATCTATGTCATCAGTGTGAGGCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTGTTGAAATGTGGATGGGTCGAACTGACAGAGGTGAGCC
GGAGTGAGGCACTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTGGAGTCAAAGAGTATGAGC
CCCAGGAAGACTGAGCAGGCCAGCAGCCCTGACTCCAACCAACATGGCCACCCAGTGACTGGTCCCCATCCT
GGGTGATGTGGCTGAAATTACACGGCTTGATAACTGTAAAGAATATTGATTACGAAGAAAACACAGCTGGAGTC
TGGGCTCTGCAATTGAGGAGTTATGAAAGAATACAATGGAAACAAACCTTTTCATCAAATCCATTGTTGAAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGTATTCTCTGTGTCATGGTAAAGTACATCAGGAATGA
TACATGCTTGTGCAAGACTGCTGAAAGAACTAAAGGAAGAATTACTCTAATCTATTGTTCTTGGCTGGCACTT
TTTATAGAATCAATGATGGTCAGAGGAAAACAGAAAATCACAAATAGGCTAAGAAGTGAACACTATATTATC
TTGTCAGTTTATATTAAAGAAAATACATTGAAAAATGTCAAGGAAAGTATGATCATCTAAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAACAAATTTAAACTACTAGTTTTTCTAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTTATATTCTATTCAATAAAAGCCCTAAACAAACTAAAATGATTGATTGTATACCCACTGAATT
CAAGCTGATTAAATTAAATTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGCCATTAAATTACAGCT
AAAATTTTAAATGCATTGCTGAGAAACGTTGCTTCATCAAACAAAGAATAATTTCAGAAGTAA

FIGURE 40

MKALLLLVLPWISPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNTRARPFERSTIRSRSFKKINR
AISVLRRTKSGAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIQHIYRDRDGVIARDGRLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAAHLIQASERRVHLVVSRQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCDKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLLAVNGRSTSGMIHACLLKGRITLTIVVSPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTCAGTTGTCATCAAGTCGAATCAGATGGAAAAGCTCAACTTGAAGCTTT
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCCATCAAATTCCGATTACTGTTGCTGTGACTTTGCGCTGACAGTGGTTGGTGGC
CACCAAGTAACACTTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTTGGGAAAGGGAAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACGTCCCTCTGTGTCCCTTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAG
CTTACAGAGGGTCGCCATCCTCGTTCCCACCGGAACAGAGAGAAAACACTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAAGTTAACGAGCCAAACTCTTGAATGTGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTAAACCTTACAAGTGTGAGGAG
CATCCAAAGCATTGGTGGTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTGG
GGGTGTTACTGCCCTAACGAGAGCAGTTTCAGGGTAATGGATTCTTAACAACACTTGAGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCCGCCCCCTG
CCTGAAGTGGTAAATATACAATGGCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAAC
GATGAAGCTTACACCAAGTGTCACTGGAGAACAGATGGTTGAGTAGTTGTTCTATAAAT
TAGTATCTGTGGAACACAATCCTTATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGTGATGTTGGAAAGAACTGATTCTTGTGCAATAATTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTATTGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCCTGGTGTGAGTATAAAACAGTTGAACAAGACAGCTTCTTAGTCATTGAT
CATGAGGGTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTTGAAGGAGATTATTAAATTGAGTAATATATTATGGGAT
AAAAGGCCACAGGAATAAGACTGCTGAATGTCTGAGAGAACCCAGAGTTGTTCTGTCAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCACTTATCCTGACAATCTGTGAAGTGGTGGTGCAGGT
GAGAAGGCGTCCACAAAGAGGGGAGAAAAGGCAGGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCCTCCAGTGTGATGCCACCAAGAGAATACATTCTATTAGT
TTTAAAGAGTTTGAAATGTTGACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAATACCTGTAGTAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN
EASTKKVELNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLIMYLLEHLHPFLQQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGVTALSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

© 2000 SAGE Publications, Inc.

FIGURE 44

MALSSQIWAACLLLLL~~L~~ASLTGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHR~~S~~KCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCACTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCGTGAAAGAGCTGGTCGGT
TCGGTTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAGCAAGTTGACTCTATTGTCTG
GACCTCAACACAACCCCTTGTCAACCATACAGCCAGAAGGGGGCACTATCATAGTACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTAAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATAACAGCTCATCACTCCAGCAGCCCTCCACCC
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATTGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGGTGTGCCAGGAACCCCTGTCAAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCAGCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCCTGTGGTGCCCTCCTGCTCAGTCTTTGACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGACTGGACATTGTCGG
AAACTCCTAACATATGCCCTATCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAACTCTAACAGGAAGATCCAGCAAATACGGTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTGTCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAASGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRRERVDFPDGGYSLKLSKLKKNDSGIYYVGIFYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSLFVLGLFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCCAAACCCCCATCTCTGCTTGAGTGGTGGTTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAAATGCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAAACAGAACTGGAATGTTCTTCATCATTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTTAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCATTGAAAAACATCAGTGACATTCCAGAATCCTCAA
CTTGCAGTGGTTTTCAATGACTCTTGGCACCTCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGGTTGGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEQFSQNPIASCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTCTCCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
CATTGCCAATAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCCTGGCCATTGCTGGGATCGCGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACGCAGCACAGTCTGTACCTGAGAAGGCCATCCACTCATCACTCCAGGCTTGCCA
CTACTTGCT**TGA**GACACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCAGCACC
TCCTCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGAAGGAGGAGCAGAGGAGGCACAGAGACGCAGAGCAAGGGCGGCAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCTGGCTGCCTCGCTGGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCAGTGGACAATATTGGGAGGCCCTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAACAGTGGCAGTGGAGTCAGGCAGGTTCCAGGCTTGGCG
AGCAGATGCTTGGCAACAGGTGGAGCAGCCCAGTGGCAGTGGAGTCAGGCAGGTTCCAGGCTTGGCG
TTGGCAGACAGGCAGAAGATGTCATTGCACACGGAGCAGATGCTGTCCCGGCTCTGGCAGGG
GTGCCTGGCACAGTGGCTGGAAACTTCTGGAGGCCATGGCATTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAAGGGCAATCTGGAGGTCTGGGACTCCGTGGTCCACGGATACCCCG
GAAACTCAGCAGGCAGCTTGGAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGG
GGGCCACAAAATTTGGGACCAACACTCAGGGAGCTGTGGCCAGGCTGGCTATGGTTCAAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGAGTGA
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGGAGCGGGAGGAAATGG
CATAAACCCGGGTGTAAAAGCCAGGGAAATGAAGCCCGGGAGCGGGGAAATCTGGGATTCA
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTCTTG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCGAGCTGGGAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAACCTCTGAGACGTCTCTGGATGTTAACCTTGACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAGGACAGAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATGGATGGAGGCCCCACACTCCCTTA
AACACCACCCCTCATCACTAATCTCAGCCCTTGCCTTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEVGKAIKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGGHIFGSQLGGQGQGNPGGLGTPWVHYPGNSAGSFGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNNEGCTNPPSGSGGGSSNSGGSGSQSGSSGSGSNGDNNNGSSSGS
SSGSSSGSSSGGGSSGGSSNSGGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHPGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGM
FNFDTWFKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAAGAGGTGTGCTGGACAAGCTGCTCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
 CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
 ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAAC TGCCGCCGGCTCCAGTGTTC
 CACAGCCCCAAAACGGAAC TGGTTTGGGTACCTGGGCTGATCACTCCTACAGAGGAGGGC
 TTGAAGGACTCGACCCAGATGTCGCCACCTATTCCCAGGGTTACGGTATGGCTGGGCTCCAT
 CATCCCCTTCATCGTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
 TTGCACCCAAGGATAATCTCTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
 AGTGGCGGTGACAAGTGGAGGCCACCGTCGGATGCTGACGCCGCCTCCATTCAACATCCT
 GAAGTCCTATATAACGATCTCAACAAGAGTGCACAAACATCATGCTTGACAAGTGGCAGCACCTGG
 CCTCAGAGGGCAGCAGTCGCTGGACATGTTGAGGACATCAGGCTCATGACCTGGACAGTCTA
 CAGAAAATGCATCTCAGTTGACAGCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
 CTTGGAGCTCAGTGCCTTGAGAGAAAAGAAGCCAGCATATCCTCAGCACATGGACTTCTGT
 ATTACCTCTCCATGACGGGGCGCTCCACAGGGCTGCCGCTGGTCATGACTCACAGAC
 GCTGTCATCCGGGAGCGCGCTCGCACCTCCCCACTCAGGTATTGATGATTTTCAAAGACAA
 AGCCAAGTCCAAGACTTGGATTCATTGATGTGCTCTGCTGAGCAAGGATGAAGATGGGAAGG
 CATTGTCA GATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACG
 GCCAGTGGCCTCCTGGTCCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
 ACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCTAAAGAGATTGAATGGGACGACCTGGCCC
 AGCTGCCCTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATC
 TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCGAGTCATCCCCAAAGGCATTACCTG
 CCTCATCGATATTATAGGGTCCATCACAACCCAACTGTTGCGGAGGATCTGAGGTCTACGACC
 CCTTCCGTTTGACCCAGAGAACAGCAAGGGAGGTACCTCTGGTTTTATTCTTCTCCGCA
 GGGCCCAGGAACTGCATGGCAGGCAGGCGCCATGGCGAGATGAAAGTGGCCTGGCGTTGAT
 GCTGCTGCACTCCGGTTCTGCCAGACCACACTGAGCCCCGCAAGGAAGCTGGAATTGATCATGC
 GC CGCAGGGCGGGCTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGCAGTGACTTCTGAC
 CCATCCACCTGTTTTGCAGATTGTCATGAATAACGGTGTGTCAAA

FIGURE 54

MSLLSLPWGLRPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRRLQCFQPPKRNWFWGHLG
LITPTEEGLKDKSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEIGILLSGGDKWSRHRMMLTPAFHFNLKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDAVIREERRTLPTQGIDDDFKDKAKSKTLDFIGVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCRCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHNPTVWPDPVEYDPFRDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGTGTTGCACTTACCTGTGTTGCCT
TTTGGTGGATAACAAGGGACTTGCACTTATCTTCTGCATTTGCAGTCTTGCAATTGACGTGG
TACAGCCTTCCTTCATACCAATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTTGC
ATAATTCATGGCCAGTTTATGAAGCTTGGAAAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACTATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTATCTGCAGCAATGTGTTGCTT
GTGATTGAACATTGAGGGTTACTTTGGAAAGCAACAATACATTCTGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGTTCTGTATCTTGAGTGGAATCTTCCTCATGTACCTGTTCCCTC
TCTGGATGTTGTCCCACTGAATTCCCAGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
AA

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGLALIFCILQSLALTWYSLSFIPFAR

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCACTGGAGCACCCAGCAGGCCAACATGCTGTCTGTGCC
TACGTGCCGGTCATCGGGAAAGCCAGACCGACTTCACTACTTTGAGTCGAAGGGCTGCCAGCTGAAGTCC
ATTTCAGCTCAGTGTCTCATCCCCTCCAGAATTCTCACCTACCGCAGTGGAAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTGAAGAATTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTGGACAAAAGAATGATGGACGATTGACGGCAGGAGATCATGCAGTCCTGCCGGACTTG
GGAGTCAGATATCTGAACAGCAGGGAGAAAATTCTCAAGAGCATGGATAAAACGGCACGTGACCACATCGACTGG
AACGAGTGGAGAGACTACCAACCTCCACCCCTGGAAAACATCCCCAGCATCTCTACTGGAAAGCATTCCACG
ATCTTGATGTGGTGAGAATTAACCGGCTGGGATGAGTTCACAGTGGAGGAGGGAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGGTGGGGCAGGGGGCTATCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCCTCCGCAGCAACACATGGGATCGTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGCCAGG
TCACCTCGCGGGGCAATGCCATCACGTCCTCAAATTGCCCCGAATGCCATCAAATTATGCCCTATGAGCAG
ATCAAGGCCCTTGTGGTAGTGACAGGAGACTCTGAGGATTACAGAGGGCTTGTGGCAGGGCTTGGCAGGGCC
ATCGCCAGAGCACATCCAAATGGAGGTCCTGAAGACCCGGATGGCCTGGCGAAAGACAGGCCAGTAACAGGA
ATGCTGGACTGCGCCAGGGAGGACTCTGGCAGAGGAGGGGGTGGCCCTTCTACAAAGGCTATGCCCCAATGCTG
GGCATATCCCTATGCCGACGTCACCTGAGCAGGCTCAAGAATGCTGGCTGAGCAGCTATGCACTG
AACAGCGCAGGGGGCGCTGTTGTGCTCTGGCTGTGGCACCTGTCAGTACCTGTGGCAGCTGGCCAGCTAC
CCCCCTGGCCCTAGTCAGGACCGGATGCAGGCCAAGGCTCTTGAAGGGCCCTGGAGGTGACCATGAGCAGCCTC
TTCAACATATCTCGGACCGAGGGGGCTTCGGCTGTACAGGGGCTGGCCCCCAACTTCATGAAGGTATCCCA
GCTGTGAGCATCACCTGCTGACGGAGACTCTGAGGAAACCTGAGATCACCTGCGAGTGGGGTGCAGTGGGGGGGG
CGCCGGCAGTGGACTCGCTGATCCTGGGGCGCAGCTGGGGTGCAGGATCTCATTTGTGAATGTGCCAACACT
AAGCTGCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGAGGGAGGGGGAGAGCTGGCAGGGCCAGGGCTT
GTCTGTGACCCAGCAGACCCCTCTGTTGGTCTCAGCGAAGACACAGGATTCCCTAGGGTCCAGGGTCA
CTCAGGCGGCTCACATGTGAAGGACAGGACATTCTGAGGCTTGTGAGGCTGGGGGGCTGGAGGGGG
TAGTCTCCATTTCACCCCTTGCGCAGGCTGTGAGGATTCTGGCTGGCCCTGGCTGTGAGGCTGGGGGG
CCTCTGCTGCCCTGTCTGCTGAGGTAAGGTGGAGGGGCTACAGCCACATCCCACCCCTCGTCCAATCCC
ATAATCCATGATGAAGGGTGGGTCACTGGCCCTCAGGCTGAGGCTTCAAGCATTGACGCCAATTGGC
TGTGAAGGAAGGAAAGGATCTGGCCTTGTGACTGGCAGGCTGAGGCTGTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAAGTGCAGGGGCTCGGGCTGGCTGGCAGGAGGCTGGAGGCTGAGGCTGGGGCTCATGGTCTGAGCT
GCCCTGGACCCCTGTCAAGGATGGGGCCACCTCAAGAACCAAACACTGTCCCCACTGTGCATGAGGGCAGTGGAGCA
CCATGTTTGAGGGGCAAGGGCAGAGCATTGTGTTCTGGGAGGGAGGGAAAGGGTGGAGGGCTTAATTATGG
ACTGGGGAAAGGGTTTGTCCAAGAGGACAAGCCGGACAATGAGCAGCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTGGCTGACTGCTCAGAGTCTGAGGCTGAGGCTGGGGGTTCTGTCCAACCCAGCAGGGCCAGC
GGGACCAGCCCCACATTCACTTGTCACTGCTGGAACCTATTATTTGATTTAATTGAACAGAGTTATGCT
AACTATTTTATAGATTGTTAATTAAATAGCTGTCAATTCAACTTCAATTTCATATTATGTTCATGGTT
GATTGTACCTTCCAAGCCCCCAGTGGATGGGAGGAGGGAGAGAAGGGGGGCTTGGGGCTGCAGTCACATCT
GTCAGAGAAATTCTTTGGGACTGGAGGAGAGAAAAGCGGCCAGAGGAGCAGGCCCTGGCTCTTCCCTGGCAG
GTTGGGGAAAGGGCTTGGCCCTAGGATTTCAGGGTTGACTGGGGGGCTGGAGAGAGAGGGAGGAACCTCAAT
AACCTTGAGGCTGAAATCCAGTTATTCCTGGCTGAGGGTTTCTTATTCATCTTCTGAAATGTCAGGCAG
TGAGGTGCTCTACTGTGAATTGTGCTGGGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCCCAGC
CTTCTGCTGCCCTGTCTAACATGCCGCAACTGGCAGCTCACGGTTGCACTTCCATTCCACAGAACCTGA
TGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAATTGTTATATGAAACATATAACTGGAGTCGCAAAAAG
CAAATTAAGAAGAATTGGACGTAGAAGTTGTCAATTAAAGCAGCCTTAATAAAAGTGTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFTSYRQWKQKIVQAGDKLDGQLDFFEFVHYLQDHEKKLRLVFKILDKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHVLVAGGGAGAVSRTCTAPLDRLKVLQMVKHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAICKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALKTGQYSGMLDCARRILAREGVAAFYKGYPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYVYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTGCCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGACATCAAACCTTGATAT
CGTGATACAATGGCTGAAGGAAGGTTTAGGCTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGCTAGAGGCCGACAGCAGTGGTGTGATCAAGTGTAGTT
GGCAATGCCCTTGGCGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGACTATAAAACTGGAGCCTCAGCATGCCGG
AACTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGGCTGAGGCTCCCGATGGTCCCC
CAGCCCACAGTGGCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAAATGGAGATCAAAGGCCGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGCT
CTCTCTTCTTGCCATCAGCTGGCACTCTGCCCTCAGCCCTTACCTGATGCTAAAATTAAT
GTGCCCTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACAGATATGACCTAGTTATTTCTGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAGAAGCCAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTCAAAGACATATTAGAAGTGGAAAATAATTGATGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCGT
CACCTGGGGAGTGAGAGGACAGGATAGTCATGTTCTGTCTGAATTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAAGCTGTAGTATGTAACCTAACGCTGCTAAATTGACTGCCACTTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGCAAATGATTCACTTTATGATGCTTCAAAGGTGCCTTGGCTTC
TCTTCCAAACTGACAAATGCCAAAGTTGAGAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAATAAAACTGAGCACCTCTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

APP_ID=10063518

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLVLHEFKEGKDELSEQDEMFRGRATAVFADQIVVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESIIRRSHLQLNSKASL
CVSSFFAISWALLPLSPYIMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTAGAATCACCATGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCTCCGGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGCTACCCCTGGTGGTTATGGGGTCCTGCCCTGGAGGGCTTATGGACCACAGCTGG
TGGAGGGCCATGGACACCCAATCTGGATGTTCCCTCTGAACCTCCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAAGCTTACGGTGCCAGCCT
GGCTTATGGACAGGGTGGCCTCCAAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGAAATTCTACCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTGTCCC
AAATGGCTACAACCTGAGCCCCCAGTCACCCAGCTGGCTCCACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GCCCTCCGGAGAAGGACACAGCTGACAAGGAAACATCCGGCTCAGCTCGAGGACTCGTCA
CCATGACAGCTCTCGGATGCTATGACCAACCATCTGTGGAGAGTGGAGTGCACCGAGGACCTT
TCCTGGCTTCTAGAGTGGAGAGAAGTATGTGGACATCTCTTCTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTCCAAAAGAGGGTGGAGAGTCTGCTGATAGCCACCA
AAATGTGAGGACCGGGCTGAGGCCACACAGATAGGGCTGATGGAGGAGAGGATAGAAGTTGA
ATGCTCTGATGCCATGAGCAGTTGAGTGGCACAGCCTGGCACAGGAGCAGGTCTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCGTGATGCCAGTGGTAGTGTCTGATCGGCTGTTACC
GTTAGTACCTGTGTTCCCTACCAGGCCATCCTGTCACACAGGCCCATTCTCCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAAAT
CCTTGTGTGTTAACTCTAGCTGCCCTGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGCAT
CTTGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCAGGAGACAGTGGCACCTCTCCCTGCCAATACTTTTTAATTGCAATTTC
ATTGGGCCAAAAGTCCAGTGAATTGTAAGCTCAATAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYGQGGAPPNVDPPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDTCLMMINMFDTKTKSGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNPQLGQPSSGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPAGGSQAVQRWPSSWGLPAMDSWPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGRTPMPHEGIWGINNQPPGTSGWNINRYPGGS
WGNINRYPGGSGWGNINRYPGGSGWNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRIQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCTTGTCCCTCTTGACCCCTTGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTTCTGACAAATTCCCTCTATGAGTCCAGCTTCTGGAA
TTGCTTAAAAGCTCTGCCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCCTCTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCGGACCCGTCTTCAGCAGGCCACCC
CTGAGTGGCAATAATAAAATTGGTATGCTG

MGSGPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSYESSFLELLEKLCLLHLPSGTSVTL

HHARSQHHVVCNT

FIGURE 66

MGSGPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSYESSFLELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATCGGTAGGAGGGCGAGCGCGAGAAGCCCC
TTCCTCGCGCTGCCAACCGCCACCCAGCCCATGGCGAACCCGGGCTGGGCTGCTCTGGCG
CTGGGCTGCCCTGCTGGCCGAGCCTGGGGCAAATACAGACCACCTCTGC
AAATGAGAATAGCACTGTTTGCCCTCATCCACCAGCTCCAGCTCCGATGCCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTGGCTGCCTGCTCTGGCTGTGGGCTG
GCACTGTTGGTGCAGGAAGCTCGGGAGAACGGCAGACGGAGGGCACCTACGGCCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATTAGGTCCCCCTCCTGCATCTGTCTCCCTCATTGCTGTGACCTTGGGAAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAaaaaaaaaaaaaaaa

FIGURE 68

MANPGLLLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSSTSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GGCAGGAATAACTAGAGAGGAACAATGGGTTATTCAAGGGTTTCTCTTAGTTCTGTGCCCTGCTGCCACAG
TCAAATCTTCTTCAATTAAAGCTGAAATAATGGCTTGAAAGATATTGTCATTGTTATAGATCTAGTGTGCCAGAA
GATGAAAAAATAATGAAACAATAGAGGATATGGTACTACAGCTTCTACGTACCTGTTGAAGCCACAGAAAAAGA
TTTTTTCAAAAATGATCTATATTAACTTCTGAGAAATGGAGGAAAATCTCTAGTACAAAAGGCCAAAACATGAA
AACCATAAACATGCTGATGTTATGGCACCCTACACTCCAGGTAGAGATGAACCATACACCAAGCAGTTACAA
GAATGTGGAGAGAAAGGCATACTTCACTTACCCCCCTGACCTTCACTTGGAAAAAAACAAAATGAATATGGACCA
CCAGGCAAACGTGTTGTCCATGAGTGGGCTCACCTCCGGTGGGAGTGTGATGAGTACAATGAAGATCAGCCTTC
TACCGTCTAAGTCAAAAAAATCGAACAGCAACAGGTGTTCCGCAGTATCTGTTAGAAATAGGTTTAAGTGT
CAAGGAGGCAGCTCTAGTAGACATGCAGAATTGAACTACAAAAGCTGATGGAAAAGATTGTCAATTCTTT
CCTGATAAGTACAAACAGAAAAGCATCATAATGTTATGCAAGATTGTTCTGTTGAAATTGAAACGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAAGTCATAATTAGTACAATGGGAGGTATTAGCAAT
TCTGAGGATTAAAACACCATACCCATGGTACACCCCTCCACCTGCTCTCATGGTCAAGAGTCAAGTCAA
AGAATTGTGCTTCTGATAAGTCTGAAAGCATGGGGGTTAGGACCCCTAAATCGAATGAATCAAGCAGCA
AAACATTCTGCTGCAGACTGTGAAAATGGATCTGGGTTGGGATGGTCACTTGTAGTACTGCCACTATTGTA
ATAAAGCTAATCCAAAATAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCTGAGG
ACTCTCATCTGCTGGAAATTAAATGCTTCAAGGTGATTGGAGGCTACATTCCAACTCGATGGATCCGAAAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCTGATGAAAGTCAAGGAAACTGGGCCCATTGTTCAT
TTTATTGCTTGGGAAAGAGCTGCTGATGAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTGTTATGTT
TCAGATGAGCTCAGAACAAATGGCCTATTGATGCTTTGGGCTCTACATCAGGAAATACTGATCTCTCCAGAAG
TCCCTCAGCTGAAAGTAAGGATTAACACTGAAATAGTAACTGCTGGATGAACGACACTGTCATAATTGATAGTACA
GTGGGAAAGGACACCTCTCTCATCACATGGAAACAGTCTGCTCCAGTATTCTCTGGGATCCAGTGGAAACA
ATAATGGAAAATTTCACAGTGGATGCAACTTCCAAAATGGCTATCTCAGTATTCCAGGAACGTGCAAAGGGGGACT
TGGGATACAATCTCAAGCCAAGGAACCCAGAACATTAAACTTACAGTAACCTCGAGCACCAAATTCTCT
GTGCTCCTAACAGTGAATGTAATGTAATGAAAGCAGTAAACAGCTTACAGGCTTACAGGCTTACAGCAGAA
ATTCTACAAAGGATATGCTTCTGAGGCAATGTGACTGCTTCTATTGAAATCACAGATGGACATACAGAAGTT
TTGGAACCTTTGGATAATGGTGCAGGCGCTGATTCTTCAAGAATGATGGAGTCACTCCAGGTATTTACAGCATAT
ACAGAAAATGGCAGATATGCTTAAAGGCTCATGGGAGGAAACACTGCCAGGCTAAAATTACGGCTTCCA
CTGAAATAGGCCGTACATACAGGCTGGTAGTGAACGGGAAATTCAAGAACCCGCCAACAGCTGAAATTGAT
GAGGATACTCAGACCCATTGGGAGATTCAAGGCCAACAGCATCGGAGGTGATTGTTGATCACAGCTTCAAGC
CTTCCCTGCTGACCAATACCCACCAAGTCAAATCACAGACCTGATGCCACAGTCACTGAGGATAAGATTATTCT
ACATGGACAGCACCGAGGATAATTGATGTTGGAAAAGTCAACGTTATCTCAAGAATAAGTGCAGTATTCT
GATCTAAGAGACAGTGTGATGCTCTCAAGTAAATACTGATCTGTCACCAAAAGGAGGCAACTCCAAGGAA
ACCTTGATTTAAACCGAAAATATCTCAGAAAGAAAATGCAACCCACATATTTATTGCAATTAAAGTATAGATAAA
AGCAATTGACATAAAAGTATCACAATTGCAAGTAACCTTGTTTATCCCTCAAGCAATCCTGATGACATTGAT
CCTACACCTACTCTACTCCACTCTACTCCACTCTGATAAAAAGTCATAATTCTGGAGTTAATATTCTACGCTGGTATTG
TCTGTTGATGGGTCTGTTGAAATTGTTAACTTATTGTTAAGTACCCATT**TGAA**ACTTAAAGGAAAGAAAATCTC
AACTAGACCTAGAAGAGAGTTTAAAAAAACAAAATGTAAGTAACAGGATATTCTGAAATTCTAAAATTCATCCCAT
GTGTTGATCAAACACTCATAAAATAATTAAAGATGTCGGAAAAGGATACTTTGATTAATGAAACACTCATGGATA
TGTAAAAATGTCAGATTAAAATTAAGTGTCTTATTGTTATTGTTAAGGAAATAGTGTGAAACAAAG
ATCTTTCTCATACTGATACCTGGTTGTATTATTGATGCAACAGTTCTGAAATGATATTCAAATGTCATAAA
GAAATTAAAATCATCTATGAGTAGTCAAATACAAGTAAGGAGGAGCAAAATAACACATTGGAAAAAA
AAA

FIGURE 70

MGLFRGFVFLLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTASTYLFE
ATEKRFKKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCAGISGRN
RVYKCQGGSCLSRAKRIDSSTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLQLTQENGWSWGMVHFSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQDGSEVLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKitGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGTLINSAWMNDVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLISPAGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNDGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRLPPLNRAAYIPGWVVNGEIEANPP
RPEIDEQTTLQDFSRATSGGAFVVSQVPSLPLPDQYPPSQTDLDAVHEDKIILTWTAPGDN
FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPDKSHNSGVNISTLVLSIGSVVI
VNFILETTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACTGCCCGGGCAGGGGTGA
 CAACAGGTGTCATCTTTGATCTCGTGTGCTGCCTTCCTAATTCAAGGAAGACGCCAAGGTAATTGACCCA
 GAGGAGCAATGATGTAGCCACCTCTAACCTTCCCTCTGAAACCCCCAGTTATGCCAGGATTACTAGAGAGTGTCA
 ACTCAACCAGAACGGGCTCCCTCGCTTAACCTGTGTTGGAGGAGAGAACCTTGTGGGGCTGCCTCTCTAGCA
 GTGCTAGAAGTGACTTGCTGAGGTGGACAGAAGAAGGAAAGGTCCCCCTTGTGTTGGCTGCACATCAGGAA
 GGCTGTGATGGAAATGAAGGTGAAACCTGGAGATTCACTTCAGTCATTGCTCTGCCGTGCAAGATCATCCTTTAA
 AGTAGAGAAGCTGCTCTGTGTTGGGTTAACCTCAAGAGGGCAGAACCTGTTAGAAGGAATGGATGCAACGAGCTC
 CGGGGCCCCAAAGCATGCTCTGTGCTAGGCCAGGGAAAGCCCTTCCGGGGGGCCCGGTTGGGGATGCC
 ACCGGTTCTGAGCAGTGGCTGATTCTGTGAATGATGTTGCGGGGGCTGCTTGGGGATTTCCCCGGTGGT
 GTTTGCTGGTCTCTGTGCTGCTATCTCTGTGCTGACATGTTGGGCTGACCCCCAAAGGTGACGAGGAGCAG
 CTGGCACTGCCAGGGCACACAGCCCCACGGGGAGGGTACCGGGCTCTCAGGGAGTGGAGGAGCAGCAC
 CGCAACTACGTGAGCAGCCAGTCAGGGCAGATGCCACAGCTCAAGGGAGCCTGAGGAGGACTGACCAAGCTCAGG
 AATGGCAGTACCAAGCCAGCGATGCTGCTGCCCTGGGCTGGAGCAGGGCCCCCAGGAAAACCCAGGGCACCTC
 CTGGCCTCTGCACTCGCAGGGACAAGGCGAGGGTAATGCTGGCGTCAAGCTGGCACAGAGTATGCAAGCTG
 CCTTTCGATAGCTTACTCTACAGAAGGTGACCGAGCTGGCTTACCCGGCACCCGGAGGAGAACCTGTG
 AGGAAGGACAAGGGGATGAGTTGGAAAGGCAATTGAATCAGGCTTGGAGGACCTGAAACATCTGAGAGAACAGC
 CCCAATCACCGTCTTACCGCCTGATTCTAGAAGGATCTACCGAACAGAAAAGGACAAGGGACATTGTAT
 GAGCTCACCTCAAAGGGGACCACAAACAGAACATTCAAACGGCTCATTTTGACGACATTCAGCCCCATCATGAAA
 GTGAAAATGAAAAGCTCAACATGGGAAACCCCTTATCAATGTTATGCTGGCTCTAGCAAAAAGGGTGGACAAGITTC
 CGGCAGTCATGAGAATTCTAGGGAGATGTCAGCAGGAGATGGGAGAGTCCATCTCACTGTTGTTACTTTGGG
 AAAGAAGAAATAATGAAGTCAAAGGAATACTTGAAAACACTTCAAAGCTGCCAACCTCAGGAACCTTACCTTCATC
 CAGCTGAATGGAAATTCTCGGGGAAAGGGACTTGATGTTGGAGCCGCTCTGGAAAGGGAAGCAACGTCTTCTC
 TTTCTGATGTTGGACATCTCAGTCATCTCAATAGCTGAGCTGATACTCAGGCAAGGAGAAGAAG
 GTATTATCAGTTCTGCTGTTGAGTCACTTCAATCCTGGCATATAATACGGCACCATGATGCTGAGTCCCTCTGGAA
 CAGCAGCTGGTCAAAAGAAGGAACCTGGATTGGAGAGACTTGGATTTGGATGACGTGTCAGTATCGGTAGAC
 TTCAATATAGTGGGTTGATCTGGACATAAAGGCTGGGAGAGATGTCACCTTATCGCAAGTATCTC
 CACAGCAACCTCATGTTGACGGCTGTGCGAGGACTCTCCACCTCTGGCATGAGAAGCGCTGATGGACGAG
 CTGACCCCCGAGCAGTACAAGATGTCAGTCAGGCAACGGCATCCACGGCCAGTGGCATGCTG
 GTGTCAGGACGAGATAGAGGCTCACCTCGCAAACAGAACAGAACAGAACAGAACAGAACATTGAACTCCCAGA
 GAAGGATGTTGGGAGACACTTTCTTCTTGTCAATTACTGAAAGTGGCAACAGAGAACAGAACACTTCCATAAA
 GGACGACAAAAGAATTGGGACTGATGGTCAAGAGATGAGAAAGCCTGGATTCTCTGTTGGGCTTTTACAACAGA
 AATCAAATCTCCCTTGCCTGCAAAGTAACCCAGTTGCTCACCCTGTGAAGTGTCTGACAAGGAGAACATGTTGTG
 AGATTATAAGCTAATGGTGTGGAGGTTTGATGTTGTTAACATACTGAGAACACTGTTGTTGTCATTG
 AATATTGATGTTAAGAGCAGTTGAAAGGACATTTGATGAGTAAAGGCAAGCATATTCTCTCATATGAATGA
 GCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAAATATCAGAAGGAGAGGAGATAGGCTTATTATGATACT
 AGTGAGTACATTAAGTAAAATAAAATGGACCAGAAAAGAAAACCATAAATATCGTGTATTTCCCAAGAT
 TAACCAAAAATACTGCTTATCTTTGGTCTCTTAAACTGTCCTGGTTTATTAACCGCTTACAGAGAACGCAACT
 TTTCTCCAGAGTAGGGACCGCTTCTACCTGTTAAAATAACCAAAGTATAACCGTGTGAAACCAAACATCTTTT
 AAAACAGGGTGTCTCCTGGCTCTGGCTCCATAAGAAGAACATGGAGAAAATATATATATATATTGT
 GAAAGATCAATCCATCTGGAGAATCTAGTGGATGGAAGTTTCTACATGTTATCCACCCAGGGCAGGTGGAG
 TAACTGAATTATTTAAATAAGCAGTCTACTCAATCACAAGATGCTCTGAAAGGAGAACACTGGAGGAAAAGAAAATGAC
 ACTTCTGCTTACAGAAAAGGAAACTATTCACTGGTGTATCGTGTACCTTAAAGTCAGAAACCCACATT
 CTCTCAGAGTAGGGACCGCTTCTACCTGTTAAAATAACCAAAGTATAACCGTGTGAAACCAAACATCTTTT
 AAAACAGGGTGTCTCCTGGCTCTGGCTCCATAAGAAGAACATGGAGAAAATATATATATATTGT
 GAAAGATCAATCCATCTGGAGAATCTAGTGGATGGAAGTTTCTACATGTTATCCACCCAGGGCAGGTGGAG
 TAACTGAATTATTTAAATAAGCAGTCTACTCAATCACAAGATGCTCTGAAAGGAGAACACTGGAGGAAAAGAAAATGAC
 ATGCATGAGCTAATTATCTCTTGAGTCCTGCTGTTGCTCACAGTAAACTCATTGTTAAAGCTTCAGAACAC
 ATTCAAGCTGTTGCTGTTAAAATGCTGATTGATTGATTGACTGGTACTGTTATGAAATTAAACACAGG
 CCATGAATGAGGTGTATTGACAGCTAATAAAATGATTGTCATGAA

FIGURE 72

MMMVRRGLLAWISRVVVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQILRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNELNMAN
TLINIVPLAKRVDKFRQFMQNFRMCMIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGNSVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
NPGIIYGHHDAPVPLEQQQLVIKKETGFWRDFGFGMTQCYRSDFINIGGFDDIKGWGGEDVHLYR
KYLHSNLIVV RTPVRGLFHLWHEKRCMDELTPEQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTCCTGGGATCCA
GAAACCCATGATAACCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCTGCTAGTCCTCTAGTCCTCAAATCCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTGTTCTCCGCCCTCTGCTGGAGGTGATTGGATCCTGGCTGCAGATGGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGCCAGCCTTACCCCTGAGTGT
GGAAACAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCCTGATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCACCCCTGATCTGGTGGACTTCCCAGAAATATGTAGCT
GCCAGCTCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACCAGATCAACAG
TGAAGCCACATTGAGAGCTCCACATTGACATTGACTCTGATTCTATGACAGCTTGAGT
AGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTGGCATCCTAATTGAGGTGGGTGAGACTAAG
AATATAGTTATGAAACACATTCTGAGTCATTGCAAGTCAGGCATAAGAGTCAGAACCTC
AGTGCCTCCCTCAACCTAACAGAGCTGCTCCCCAAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCAGAGTGTGCTCTGGACAGTTTTATAGAAGGTCCCAG
ATTCAATGAAACAGCTGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAAGAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGATGGCTTGCTTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTAAATGCTGAGTCTAGGTGAGGATCTGGTGG
TGCTCTGCCTCTCCTGGCTTTATTCTATTGCTAGAAAGATTGGAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGTCTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCTCTCAGATAC
CATGGATGTGGATGACTTCCCTTCATGCCCTACAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCCTGGACATCTCTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGA
ATGAGGAAATCGCTGTGTTAATGCAGAGANAAACTCTGTTAGTTGCAGGGGAAGTTGG
ATATACCCCAAAGTCCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCAGGGATCT
CTCCTAGGATAAAGAGTTGCTGTTGAAGTTGATATTTGATCAATATATTGAAATTAAAG
TTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFDPPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGGCGCTGCTGCTGTTGCCTGGCGCGCCTGGGGACGGGCAGTCCCTGTGTC
TCTGGTGGTTGCCTAACCTGCAAACATCACCTCTTATCCATCAACATGAAGAATGCCTACA
ATGGACTCCACCAGAGGGCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGCTTGACAGCTCC
AGAGAAGTGGAGAGAAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCAGTGTGACCAACCAC
ACGGTGGTGCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTGCCATATCTATTACCGTGTTCCTTTT
TCTGTGATGGGTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
GATTTGATTTATGAAATGAATTGACAAAAGATTCTTGTGCCGTGAAAAAAATCGTGATTA
ACTTTACCCCTCAATATCTCGGATGATTCTAAATTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGTGATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGTATGCTCGCATTGATGAAATTGGACTCTGAAG
AAAACACGGAAGGTACTCTCTCACCCAGCAAGAGTCCTCAGCAGAACAAACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAAGAACACTGACATTGTCGGGGCTGAGAGCAGGA
GCTCAGTTGCAGGAGGGGTCCACACAAGGAACATTGGAGTCGCAGGCAGCCTGGCAG
TCTTGGGCCGAAACGTTACAGTACTCATACACCCCTCAGCTCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGTCAGTGGGA
TCCCCAAACTGGCAGGCTGTATTCCCTCGCTGTCAGCTCGACCAGGATTGAGGGCTGCG
AGCCTCTGAGGGGATGGGCTCGGAGAGGGCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAAGGAGAAAATGAAACCTATCTCATGCAATTCTAGGAGGAATGGGGTTATA
TGTGAGATGGAAAACTGATGCCAACACTCCTTTGCCCTTGTGCAAAACAGTGAG
TCACCCCTTGATCCCAGCCATAAAGTACCTGGATGAAAGAAGTTTCCAGTTGTCAGTGT
CTGTGAGAATTACTTATTCTTTCTATTCTCATAGCACGTGTGATTGGTTCATGCATGTA
GGTCTCTAACATGATGGTGGGCCTGGAGTCCAGGGCTGGCGGTTCTATGCAGAGAA
AGCAGTCATAAAATGTTGCCAGACTGGGTGCAGAATTATTACAGGTGGGTGT

FIGURE 76

DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIGNEFDKRFVPAEKIVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEVKHTMLGYASHLMEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTDICAGPEEQELSLOEEVSTQGTLLESQAALAVLGQPQLQYSYTPQLQLDPLAQEHTDSEEGPPEEPSTTLVDWDPQTGRLCIPSILSSFDQDS

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTCCCCAGGTCTGGCATCTGCAC TTGCTGCCCTTGACAC
CTGGGAAGATGGCCGGCCCCGTGGACCTTCACCCCTCTGTGGTTGCTGGCAGCCACCTTGATC
CAAGCCACCCCTCAGTCCCCTCAGTCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTAACACCGCTCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTCCACATGACCACTGAGGCTGAGGCCAAGCCACCATCCGCATGGACACCACTGCAAGTGGC
CCCACCCGCCTGGTCCCTCAGTGCACCTGGGAGCCTGCGCATCCAACTGCTGTA
TAAGCTCTCCTCCTGGTGAACGCCCTAGCTAACAGCAGGTCAAGACCTCCTAGTGCCATCCCTGC
CCAATCTAGTAAAAACCAAGCTGTGCTCCGTGATCGAGGCTTCCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTGAACACCATTAGCTCTACCTGGGGGCCAAGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTCCCTGACAATGCCACCCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTAAAGCTGCAGTGGCTGCTGTGCTCTCCAGAAGA
ATTCAATGGTCTGTTGGACTCTGTGCTCCCTGAGAGTGCCATCGCTGAAGATCCTAACCTCAGGAC
TGATCAATGAAAAGGCTGCAGATAAGCTGGATCTACCCAGATCGTAAGATCCTAACCTCAGGAC
ACTCCCGAGTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGTGGAAGTGTT
TCCCTCCAGTGAAGGCCCTCCGCCCTTGTGACCGTCTGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAGGTGACCAACTTAACTCAACTGAAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGATTGGCTGGTCCAACTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGCCAATTAGATCTGGGCTCCAGTGTGTTACTCCAGCCTCC
CCTTGGGATTGCCAGCTCTCTGAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCTGAGCTTCCAGTGAAGACTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCAGCTGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELDHNATSILQQLPLLSAMREK
PAGGI PVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSRLIQQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPTLDNIPFSLIVSQDVKAAVAAVLSPEEFMVLLDSVLPEAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQIMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCACCCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGCCTCT
CTTGGCCTCCAACCTGTGGGCTACATCCTAGGCCTCTGGGGCTTTGGGCACACTGGTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTCTTATGTCGGTGCACAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTCTGGGCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGGCTGCATTATCTCTGGTGGCATGAGATGACAGTCTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTCTGGGATTC
ATT CCT GTG C CT GGA AT C TT CAT GGG AT CCT AC GGG ACT T C T A C C ACC ACT GGT G C CT G AC AG
CATGAAATTGAGATTGGAGAGGCTTTACTTGGCATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTCTGCTCATCCCAGAGAAATCGCTCAAACACTACGATGCC
CAAGCCCAACCTCTGCCACAAGGAGCTCTCAAAGGCCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAGAACCCAGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAACAGTGGACAGCACCCCGAGGGCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCATTGGATTGAGCAAAGGCAGAAATGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTTGCTGCTCCCTGCCCTAAGTCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCCTCTGTTTACCTGGACTCCATCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTCTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCACTTTGTTATGACTCCACAGTGTCCA
GACTAATTGTCATGAACGTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTAAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTIVAMILLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSIASSILACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPIVAWNLHGIILRDFYSPLVPDSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACCGCGTCCGGCCTCTCCCTCTGCTGGACCTCCTCGTCTCTCCATCTCTCCCTCTTC
 CCCCGCTTCTCTTCCACCTTCTCTTCTCCCACCTTAGACCTCCCTGCCCTCCCT
 GCCCACCGCTGCTCCCTGGCCCTCTCCGACCCGCTAGCAGCAGACCTCTGGGGCTGTGG
 GTTGAATCTGTGGCCCCCTGTGCCTCCGTGCTCTTCGTCCTCCCTCCGACTCCGCTCCGG
 ACCAGCGGCCTGACCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCCCTCCCTGCTGGGA
 CTCGCCTGCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
 CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
 ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCATGTGAGTTACCGCCCTCCACTGTCGCCT
 GTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGCTCCAAGTGTGGAACCTCACAC
 TCCCTCTGGACTCCGGGCCACCAAAGTCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCCCTGCCAACCAAGTGTGTCCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCCTCACAAACCTGCCCGAACCCAGGCTGCCAGCACCCCTCC
 ACTGCCAGACTCCTGCTGCAAGCCTGCAAAGATGAGGCAAGTGTGAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGAGAAAG
 AGAGGCCCGGGACCCCAAGCCCCACTGGCCTCAGGCCCTCTGAGCTTCCATCCCTGCCACTT
 CAGACCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGCCAGGACTGCCAGCTGTGACCTGTCC
 GTGTGCATGGGGAGACGTACTCCCACGGGAGGTGTGGCACCCGGCTCCGTGCCCTCGC
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCAGGACTGCCAGCTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTACCCCGAGAAAAGTGGCTGGGAAGTGTGCTGCAAGATTTGCCAGAGG
 ACAAAAGCAGACCCCTGCCACAGTGAGATCAGTTCTACAGGTGCTCCAAAGGCACCGGGGGTC
 CTCGTCCACACATCGTATCCCAAGCCCAGACAACCTGCGTCGCTTGTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGAAAAGATGAGGAAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCAAGGCCACACAGCCAGAATCTTCACTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCACTGCTGCTGGCCCCACGAAGGTCACT
 GGAACGTCTTCTAGCCAGACCCCTGGAGCTGAAGGTACGGCCAGTCCAGACAAAGTGAACCAAG
 ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATATTAAATAAA
 TAAGAAGTTGCATTACCCCTAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTTCPPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGT PAPTGLSAPISFIPRHFRPKGAGSTTVKIVLKEHKKACVHGGKTY
HGEVWHPAFRAFGPLPCIILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARI PERGTALPTARWPPRRSLERLPSPDPGAEGHGQSRQSDQDITTKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCGATGGAGTAGACTCTAGAACAGGCCAGTTGCCCTCCGCTCACCGCAGAGCCTCTCC
 GTGGCTTCCGCACCTTGAGCATTAGGCAGTTCTCCTCTAATCCATCCGTACACCTCTGTCA
 TCCGTTCCATGCCGTGAGGTCATTACAGAACACATCCATGGCTCATGCTCAGTTGGTTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGCCAGACAAGCCTGTCCAGGCCTGGTGGGGAG
 GACGCAGCATTCTCCTGTTCTGTCCTAAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACAGCCATTATGCAGATGCCACAGTATC
 AAGGCAGGACAAACTGGTGAAGGATTCTATTGGGAGGGCGCATCTCTGAGGCTGGAAAACATTACT
 GTGTTGGATGCTGGCCTATGGGTGCAGGATTAGTTCCACTTACTACCAGAACGGCATCTGGAGCT
 ACAGGTGTCAGCACTGGGCTCAGTTCTCTCATTCACAGGGATATGTTGATAAGACATCCAGCTAC
 TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTGTCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCTGTTGATGTTGAGATCTCTGACCGTCAAGAGAA
 CGCCGGGAGCATACTCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
 GAGATACCTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTTGGCATTGTTGGACTGAAGATTTCTCTCAAATTCCAGTGGAAAATCCAGGCGGAACTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCCAACAGCGAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTACCCGAAGCTCTGCGTTCTGATCTGAAAACCTGTAACCCATAGAAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
 CTGGGAGGTGGACGGAGGACACAATAAAGGTGGCGCTGGGAGTGTGCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCCTCCGATCATGGTACTGGGCTCTCAGACTGAATGGAGAACATTGTATTTC
 ACATTAATCCCCGTTTATCAGCGTCTCCCAGGACCCACCTACAAAAAATAGGGTCTCTGGACTA
 TGAGTGTGGGACCATCTCCTCTCAACATAATGACCGACTCCCTATTATACCCCTGACATGTCGGTTG
 AAGCCTATTGAGGCCATCATTGAGTATCCGCTCTAATGAGCAAATGGAACCTCCATAGTCATCTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTTGGCAAAGGGCTCTGCAATCCAGAGACAAGCAACAG
 TGAGTCCTCTCACAGGCAACCACGCCCTCCCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
 TCTCTTAGGGATATTAGGTCTCTCCTCCAGATCCAAAGTCCCGAGCAGCCGGCAAGGTGGCTTCCA
 GATGAAGGGGGACTGGCCTGTCACATGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGAGGGAAAGAAGG
 CTGACATTACATTAGTTGCTCTCACTCCATCTGCTAAGTGAATCTGAAATACCACCTCTCAGGTGAAG
 AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTACAGAACAGGAATGTGAATAATGCTTAG
 ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTTCAATTACACTTCAGTAAAAAA

FIGURE 84

MALMLSLVSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVP LISITGYVDRDIQLLCQSSGWFPRPTAKWKGPGQDLSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKPQEVPHSEKRF
TRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPPPTKIGVFLDYECGTISFFNINDQSLIYTLCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCCCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGCCCACTAGTTCATGGCTACTGGTTCGGGAAGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGCGGGGAGA
TACCTCTTCGTATGGAGAAAGGAAGTATAAAATGAAATTATAAACATCACGGCTCTGTGAA
TGTGACAGCCTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCCCT
AGAATCTGACCTGCTGTGCCCTGGGCCCTGTGAGCAGGGACACCCCTATGATCTCTGGATA
GGGACCTCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCACA
GCCCCAGGACCATGGCACCGCCTCACCTGTCAGGTGACCTCCCTGGGCCAGCGTACCGACGA
ACAAGACCGTCCATCTCAACGTGTCCTACCCGCTCAGAACTTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGAAATGGCTCATCTGTCACTCCAGGGCCAGTCT
GCGCCTGGTCTGTGCAAGTTGATGCAAGTGACAGCAATCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGTGGAGCTGCCCTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAAGAGCTCAGAACCCCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTTCCTGCGTCATCTCGTTGAGTGAGGCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATAACGGGCATAGAGGATGCAAACGCTGTCAGGGGTT
AGCCTCTCAGGGCCCTGACTGAACCTTGGCAGAACAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCGCTCCTCAGTGGGGAGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGATGAGCTATGATAAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAAGACAGACAAATTCTA

FIGURE 86

MLL LLL PLLWGRERAEGQTSKLLT MQSSVT VQEGLCVH VPCSF SYP SHG WI YPGPVV HGY W FREG
ANT DQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTL SIRDARRSDAGRY FFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPOPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLNGSSL
SLPEGQSLRLVCAVDADSNNPPARLSLSWRGLTLCPSQPSNPVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKK SARPAAGVGDTGIE
DANAVRG SASQGPLTEPWAEDSPPDQPPPASARSSVGEGELOQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTCCTGCTGTTTC
TCATAGCGACCACCAGAGGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCAGAACGCTGCAAGGAATCAAAGACGAATGCTCTAGTGCAATTG
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCACAAAGCAGACTACCCAGAGGGGACGGCAACTGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCTCCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTTCTATGATTTGGCAGCAGCCAGAAAACAGCATCTTATTACTCACCCCTAT
GGCAGCGGGAAATTCACTGCGGGATTTGTCAGTTCAAGGGTATTAATAACGAGAGAGCAGCCAA
CGCCTTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACATGTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAAACCCAGACCTCTCCCAACCATGAGATCCAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPFGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVI PVVYDFGDAQKTASYSPYQREFTAGFVQFRVFNNERAANALCAGMVRTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGTTGC~~GGGG~~GACTCAGGAGTCGCTGTCTGAAC~~TCCAGCCT~~CAGAGAC
CGCCGCC~~TTG~~CCCCGAGGGCC~~ATGG~~CCGGGTCTCAGGGTTGTGCCCTCGCTTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCC~~GGG~~ACAGCAACATACAGGCCTG
CCTGCC~~CT~~TCACGTTACCCCCGAGGA~~G~~TATGACAAGCAGGACATT~~CAG~~CTGGTGGCCCGCTCT
CTGTCACCCTGGCCTCTTG~~CAGT~~GGAGCTGCCGGTTCC~~CT~~TCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCAT~~CTCC~~ATTGGG~~G~~CTCACTGTAGTGCATCCGTGCC~~CT~~GTCC~~TT~~CCAT
ATT~~CGAGCG~~TGGGAGTGC~~ACT~~ACGTATTGGTACATT~~TT~~GTCTTCTGCAGTGCC~~CT~~TCAGCTG
TC~~ACT~~GAAATGGCTTATT~~CGT~~CACCGTCTTGGG~~G~~TAAAAAGAAACCC~~TT~~CTG~~ATT~~ACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTCGTATTCC~~GT~~GAAGAGGAAG
GCATAGGCTT~~CGG~~TTTCC~~CTCG~~AAACTGCTTG~~GAGG~~ATATGTGTT~~GG~~AATAATTACG
TCTTGAGTCTGGATTATCGCATTGTATTAGTGC~~TT~~GTAA~~AAA~~ATGTT~~GT~~AGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAA

...MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKKPF

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCC GAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAAGAATGCAACTGAC
TCGCTGCTGCTCGTGTCCCTGGTGAGGGTAGCCTATCTGGCATCTGTGCCAGGATGATG
GTCCTCCCGGCTCAGAGGACCTGAGCGTGTGACACAGAGGGCAGCCCCGGCCCGGGTGCCT
CGGAAGCGGGGCACATCTCACCTAACGAGGCCCCATGGCAAATCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGAGGCTTGGGATTCTTGGCAGGCCCCAACGCCCCAACGACCACAGCCCC
CACCCCTAGCCAAGGTGAAGAAAATCTTGGCTGGGCACTTCACTCCAAACATCAAGACGGTG
GCCCTGAACCTGCTCGTACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCTCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCGGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGGCCCTCAAAGTCGTCTGTG
TCTACATCGCCTTCACTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACACT
CATAGTATAACCCCTACTACCCATCTGGGTGACCCGGGCAGGCCACAGAGGCCAGGGC
TGGAAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCAGGGCAGGGAGGGTTGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGGAGATGCCAAGTGGGCAAGGCTCAAGTGGCAG
AGAAAGGGTCCAAGTGTGGTCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGG
TCCCCGAGGCCTGTGGGAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAGC
CCTTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCCTAGGAGCCAGTCAGCAGGGTGGGTGG
GCCAGAGGAGCTCCAGCCCTGCCAGTGGGCGCCCTGAGCCCTGTGCTGTGAGCATGG
CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGTCTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATTCCCTCTGCCAGTACTCCCCCTGTACCAACCCATTGCTGATG
GCACACCCATCCTTAAGCTAACAGACAGGACGATTGTGGTCCCTCCACACTAACGCCACAGGCCATC
CGCGTGTGTGTCCCTCTCCACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTAACAGTCAGCCTCACCTGTCAGACCCGGGTTCTCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGTGGGCGGGCCGGCCGAGAGCATGTGCTGGATCTGTC
TGTGTGTCTGTGTGGGTGGGGAGGGAGGGAGGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTGGAGCAGGAATAAGCTTGGCCGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMA
NSTLGLLAPPGEAWGILGQPPNRPNHSPPPSAVKKIFGWGDFYSNIKTVALNLVTGKIVDHGN
GTF SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRR
TSLC THDP AKICSRDHAQSSATWSCSQFKVVVCVYIAFYSTDYRLVQKVCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGC~~GGCCGT~~TCTCGGCTGCCCTTCATTGCC~~TT~~CGGGCTGCCTCGCC
CTTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATT~~CGT~~CCCTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCTCTGTCTATATCCAAGAA
ATGTTCCGATTGCATATTATAAACTCTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGCCATGTTCTGGCTGGGCTTGGAAATCATGA
GTGGAGTATTTCCTTGTGAATACCCATCTGACTCCTGGGCCAGGCACAGTGGCATT~~CAT~~
GGAGATTCTCCTCAATTCTCCTTATT~~CAG~~CTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGCATTGATTTTGATGGCTGTGAGAAGAAAAGGGCATCCTCCTATCGTTC
TCCTGACCCACCTGCTGGTGT~~CAG~~CCCAGACCTTCATAAGTTCTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGT~~CAT~~GGGCACCTGGCATTCTAGCTGC~~GGGAGGCAG~~CTGCCG
AAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACTTCTCTTACAACCAGCGCTCCAGA
AACCTCAGGAACCAGCAC~~TT~~CCAAACCGCAGACTACATCTTAGAGGAAGCACA~~ACT~~GTGCCT
TTTCTGAAAATCCCTTTCTGGTGGATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLIFGAFVSYYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGP GTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKGILLIVLLT
HLLVSAQTFISSLYYGTINLASAFIILVLMGTWAFLAAGGSCRSLKLCQDKNFLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTGAGAAACCAACTGGACCTTGACTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGCTTCTAGGATCAAC
TCGGTCATTACCACAGCTAAACCTGCTTGGACTCCCTCCCACAAAATGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTCTTAAGTCTGATACCATTAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCCAGCAGGCCCTCCAACCTCCAGGGCACAGATGACGACTTGCAGTGACCAACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGAAATTCA
GTAAGCTGTTCAAATTTCAACTAAGCTGCCTCGAATTGGTGTACATGTGAATCTTATC
ATTGATTATAATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAAATATTCTGAAATTTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTAAAAAA
CAATAATTCAATGGATAAATCTGTCTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAATGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLPDPLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGCTCTCCGCCCCAGGCTCTCCCAAGGCCAGCCTGC
 AGAGCTGTCTGTGGAAGTCCAGAAAATATGGTGAAGATTTCCCTTATACCTGACCAAGTTGC
 CGCTGCCCCGTGAGGGGCTGAAGGCCAGATCGTGTCAAGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTGCTATGGATCCAGATTCTGGCTTCCGTGGTACCAGGGCCCTGGACCAGGGCCCTGGACCAGAGGA
 GCAGGCAGAGTACAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGCTTGTGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCTTCCTCTGAGGCTTCAGACCG
 AGAGCTCGGCTGAGCCGGGTACCAAGGCCCTGGCATCCCTTCCTGAGGCTTCAGACCG
 GGATGAGCCAGGCACAGCAACTCGGATCTCGATTCACATCTGAGGCCAGGCTCCAGGCCAGC
 CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCAACGCCCTGGAGGACCTACAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCCCTCAGGCCACCAGGCCACTGCCACCGTGAAGTCTCATCATAGAGAGCACCTGGGTGT
 CCTAGAGCTATCCACCTGGCAGAGAATCTAAAGTCTATAACCCGACACATGGCCAGGTA
 CACTGGAGTGGGGTATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCTTGAAGTGAA
 TGCAGAGGGAAACCTCTACGTGACCAAGAGAGCTGGACAGAGAACGCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTCCCATGGCAGGAGACTATGCGGCCCCCTGGAGCTGCACGTGCTGGT
 ATGATGAGAATGACAACGTGCATCTGCCCTCCCGTGACCCACAGTCAGCATCCCTGAGCT
 CAGTCCACAGGACTGAAGTGAAGTCACTGTCAGCAGAGGATGCAGATGCCCGCTCCCCCA
 ATTCCACGTTGTGATCAGCTCTGGCAGGACTATGCGGCCCCCTGGAGCTCCAGTCTGAGCT
 CAGGGTGGACCCACTTCAGGCAGTGTGACGCTGGGGTGCTCCACTCGAGCAGGCCAGAACAT
 CCTGCTTCTGGCTGGCATGGACCTGGCAGGCCAGAGGGTGGCTCAGCAGCACGTGTAAG
 TCGAAGTGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATTGGCCT
 ATAAGCCTCCCTGAGGATGTGGAGGCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCCTTCCGCTCATGGATTTGCCATTGAGAGGGAGACACAGAACGGACTTTG
 GCCTGGATTGGAGGCCAGACTCTGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCACTGAGGGTGGTGGTGACAGTGTGCTAGTGGAGAGACTGATGCCACCCCCCAAGTTGG
 AGGGCCCTGGAGGCCACGGCAGCTGACTGTCAGTGGCAGAGGTGGCAAGCTGGTGGGGCAGGCC
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGGCCAACCTCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATTCTCCGGGGAGGTGCACACCGCCAGTCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCTGTGCCCTCCAA
 TACCTCTGCACACCCGCCAACGACCATGGCTGATCGTAGTGGACCCAGCAAGGACCCGATCT
 GCCAGTGGGCACGGTCCCTACAGCTTCACCCCTGGTCCCAACCCACGGTCAACGGGATTGGC
 GCCTCCAGACTCTCAATGGTCCCCATGCCCTACCTCACCTGGGCTGCACTGGTGGAGGCCAGT
 GAACACATAATCCCCTGGTGGTGAGCCACAATGCCAGATGTGGCAGCTCTGGTGGAGCT
 CGTGTGCGCTGCAACGTGGAGGGCAGTGCATGCCAAGGTGGGCCAGTGAAGGGCATGCCA
 CGAAGCTGTCGGCAGTGGCATTCTGTAGGCACCTGGTAGCAATAGGAATCTCTCATCCTC
 ATTTCAACCAACTGGACCATGTCAGGAAGAAGGACCCGGATCAACCAGCAGCACAGCGTCCCCT
 GAAGGGGACTGTC**TGA**ATGGCCAGGCAGCTAGCTGGAGCTGGCTCTGGCTCATCTGAG
 TCCCTGGGAGAGAGGCCAGCACCAAGATCCAGCAGGGACAGGACAGAGTAGAAGCCCTCCA
 TCTGCCCTGGGGTGGAGGCACCATCACCACCAACGGCATGTCAGAGGCCAGACACCAACT
 TATGGACTGCCATGGAGTGCTCAAATGTCAGGGTGTGCCCCATAATAAGCCCCAGAGAA
 CTGGCTGGCCCTATGGGAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAELYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPD MFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLPEIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLQVRAQNSHGEDYAAPPLEHVL
VMDENDNVPICPPRDPVT S IPELSPPGTEVTRLSAEDADAPGS PNSHV VYQLLSPEPEDGVGRA
FQVDPTSGSVTLGVPLRAGQNILLV LAMDLAGAEGGF SSTCEVEVA VTDINDHAPEFITSQIG
PISL PEDVEPGTLVAMLTAIDADLEPAFR LMDFAIERGDTEGT FGLDWE PD SGHVR LRLCKNLSY
EAAPSHEVVVVVQSVAKL VGP GP GP GATATVTVLVERVM PPPKLDQESYEASVP ISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWL CIEKFSGEVHTAQS LQGAQ PGDTY TVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGP NPTVQR DWRLQ TLNGSHAYLT LALHW VEP
REHIIIPVVVSHNAQM WQLL VRVIVCRCNVEGQCMRKVGRMKG MPTKLSAVGILVGT LVAIGIFLI
LIFT HWTMSRK KDPD QP ADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTA
TCTCTTCACCTCAAGTCCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATTTGCTTCAGCCACAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCTTATGTTGGTCACTATTGCAATTAGAACGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCAGTGGAGCCAGCACAGCCACCAACTCTGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAAACCTCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAA
CAACTCTGACTCCAGCACAAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAAACCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCAAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCAGTGGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTAGGGCCAAACAGCCACCAACTCTGAGTCCAGTGGGCCAGCAGTGGCCACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCAGTGGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCAG
GCCACCAAACCTGACTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAAACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGCAAGCTGAGTCCAGCACAGCAGCTCTGACTGGA
ATGACACACAATTCCCATAGTCATCTACTGCACTGAGTGGAGCAAAGCCTGGTGGTCCCTGGT
GCCGTGGAAAATCTCCCATCACCCTGGTCTGGGCTGGCCGGCTGGGCTCTTGTGGC
TCTCTTCTGTGTGAGAAACAGCCTGCTCTGAGAAACACCTTAACACAGCTGTCTACCAACCT
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCTTAAGTGGTTCTGGAGGGAGACCAGTATCATCGATAGCCATGGAGATGAGGGGAGGA
ACAGCGGGCCCTAGCAGCCCCGGAAAGCAAGTGGCGATTCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTTCATTCATCCCAGGAGACCCCTCCAGCTTGTGAGATCCTGAA
AATCTTGAAGAAGGTATTCTCACCTTCTGCTTACAGACACTGGAAAGAGAAATACTATAT
TGCTCATTAGCTAAGAAATAATACATCTCATCTAACACACAGCACAAAGAGAAAGCTGTGCTTG
CCCCGGGGTGGTATCTAGCTCTGAGATGAACCTCAGTTATAGGGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTTSEASSTATNSESSTVSSGISTVTNSES
STTSSGANTATNMGSSVTSGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSSI
AMEMSGRNNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCCCGCTTACGGGATGAATTAAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCTGCTTGTACAGGTG
GGAGGCTGGAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATA**TG**CATTCCAGGGAAATTGTGGTGCCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTCATAAAGCAGGACTCTAAAAGCTTGGAAATCAT
GGTGTCACTGAAAGGGATTACTTACTGACTCTGTTGGGAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTGTGGCAACATGGCTCACCCACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCACCGAACAA
GAATGGACTGGATGTTCTGTGGATTGCCTGATGCGATATACTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCAAGGAAATGAAAGGATGACAAGAGCCTTCGAAGACATGATTGATTACTTTGT
ATATTCAACGAACCACTTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCAAGAGAAGGTAAAGAACCTGATGCTGTCATG
ATATCACTGTGGCTATCCTCACACATTCTCAATCAGAGAACGACCTCCAAAGGAGACTTT
CCCAGGGAAATCCACTTCACTGTCACGCCACCGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCACAAACGGTGGGAAGAGAAAAGAAGAGAGGGCTGCCCTCTATCAAG
GGGAGAAGAATTTTATTTACCGGACAGAGTGTCAATTCCACCTTGAAGTCTGAACCTAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCCTGTTAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTGTAAAGTGGTATTTATAATCACCATTGTAATCTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACCTGCAATTGACTTACGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAG**TAAG**ATTAAAGGTTGCCATGTGAAAACCTAGAGCATATTTG
GAAATGTTCTAAACCTTCTAACGCTCAGATGCATTGCACTGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAAATTGTTGGAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGCAACTGTGTACATAGCAGGGAGTGATGGGTGAAATAACTT
GGGCCAGAATATTAAACAAATCATCAGGCTTTAAA

... EBB CCB CCC CCB AAB CCC ... m GGG GCG GCG GCG GCG GCG

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMWSKGIIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLGVPFGWAMQAAAYIFIHRWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDNAVHDITVA
YPHNIPQSEKHLLQGDFPREIHFHVRYPIDTLPTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGACTGAAGAGCCTCTCCACGGCTCCTGCCCTGAGACAGCTGGCCTGACC
TCCAATCATCCATCCACCCCTGCTGTATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTTGGTCTCAGTTCTACGAGCTGGTGTCAAGGACAGTGGCAAGT
CACTGGACGGCAAGTTGTCCAGGCCCTGGTGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
TTCCTGAGACCACTGCAGAGGCTATGAAAGTGCAGGGTCTCAGGAATCAGTCCATGCTGGTC
CACCTCTACAGAGATGGGAAGACTGGAATCTAACGAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGAGGACTCCATTGCAAGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATGGCCTGTATGGGTGCTGGTCAGTCCAGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTACTTCCCTCTCATTTCCATCGTGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCCTGTCTCAGGCTGGTCCCCCAGCCCACAGCCAAAGGATGCCAAGGAC
AGGATTGCTTCAGACTCCAGAACATGCAAGATGGGTACAGCTGTATGATGAGGATCTCC
ATTATAGTCCAGGAAATGCTGGAGCATATTGTTCCATCCACCTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTGGCGCCTGGCTCTA
TTTACTCGGGTTACTCTGTTGCCCCGTGTGGTGTGATGGGATGATAATTGTTCTC
AAATCAAAGGAAAATCAGCGGAACCTGGACTGGAGAAGAACGACGGACAGGCAGAATTGAG
AGACGCCCGAAACACGCACTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGCG
TTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCAGGGAGGTGCTCACTCTGAGAACAGA
TTTACAAGGAAGAGTGTGGCTTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAAGGGTGTGGAGTGTGGAGCTGAGCTGAGACAGGGGAAAGAAC
ATGTGACTTTGCTCCCAACATGGTATTGGGTCTGGGTCTCAGACTGACAACAGAACATTGATTTC
ACATTCAATCCCCATTTCATCAGCTCCCCCAGCACCCCTCTACACGAGTAGGGTCTTCCT
GGACTATGAGGGTGGGACCATCTCTTCTCAATACAAATGACCAGTCCCTATTATACCCCTGC
TGACATGTCAGTTGAAGGCTGTTGAGACCCTATATCCAGCATGCGATGTGACGAGGAAAG
GGGACTCCCATATTCAATATGTCAGTGTCTGGGATGAGACAGAGAACCCCTGCTAAAGGGC
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTCCTCT
CCGGAGCCTGCGCACAGAGACTCACGCCCTTCTAGGGAGCTGAGGTTCTGCCCC
TGAGCCCTGCAAGCAGGGCAGTCACAGCTCCAGATGAGGGGAGATTGGCTGACCCCTGTGGGAG
TCAGAACGCATGGCTGCCCTGAAGTGGGAGGAATAGACTCACATTAGTTAGTTGAAAA
CTCCATCCAGCTAACGCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG
ACAGTGATTCTGCTCACAGGTGAAGGATTAAGAGAACAGAACATGTGAATCATGCTTCAGGTT
TGAGGGCACAGTGTGTTGTAATGATGTTTTATATTACATTTCCACCATAAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTATACCAAATACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTCAAAGATAAAGAGGAGGTAGGATTTCACTGATTCTATAAGCCAGCAT
TACCTGATACCAAAACCAGGCAAAGAAAACAGAACAGAACAGAGGAAGGAAACTACAGGTCCATATCC
CTCATTAACACAGACACAAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATATTAA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAAGGTTAATATTAAAT
ATCAACCAGTGTAAATTCAACATTAATAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 104

FIGURE 105

CCTTCACAGGACTTTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
 TTTGGTGGGAACCTGGTTATCGGCCTCGCATCTCATCCTGATTGTCTGGCAGTGTGCATTGGA
 CTCACTGTTCAATTATGTGAGAGATAATCAAAGAACCTACAATTACTATAGCACATTGTCAATTACAAC
 TGACAAACTATATGTGAGTTGGCAGAGAGGCTCTAACAAATTACAGAAATGAGCCAGAGACTTGAAT
 CAATGGTAAAAATGCATTTATAAATCTCCATTAAGGAAGAATTGTCAAGTCTCAGGTTATCAAGTTC
 AGTCAACAGAACGATGGAGTGTGGCTCATATGCTGTTGATTTGAGATTCACTCTACTGAGGATCCTGA
 AACGTAGATAAAATTGTTCAACTTGTACATGAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
 ATCCTCACTCAGTAAAAATTAAAAAAATCAACAAGACAGAACAGACAGCTATCTAAACCATTGCTGCGGA
 ACACGAAGAAGTAAACTCTAGGTCAAGACTCAGGATCGTGGGAGCAGAAGTAGAAGAGGGTGAATG
 GCCCTGGCAGGCTAGCCTGCAGTGGGATGGAGTCATCGCTGTGGAGCAACCTTAATTATGCCACATGGC
 TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCCCTGCGAGATGGACTGCTTCCTTGGAGTAACA
 ATAAAACCTTCGAAATGAAACGGGGTCTCGGAGAATAATTGTCATGAAAAATACAAACACCCATCACA
 TGACTATGATATTCTCTTGAGAGCTTCTAGCCCTGTTCCCTACACAAATGCACTAGATAGAGTTGTC
 TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGACAGGATTGGAGCACTGAAAAAT
 GATGGTTACAGTCAAATCATTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
 TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
 GCCAGGGTGACTIONGGAGGACACTGGTTAGTTCACTGAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
 AGCTGGGAGATGAATGCGAACCCAAACAAGCCTGGTGTAACTAGAGTTACGGCCTGGGGACTG
 GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTGTTGTTTG
 GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
 ATAAACTGTTGTTGATGCACTGTTGAGCAATAGTTGAAACATTATGTACATAGAGAAATAGATAATAAATATTAC
 ATTACAGCCTGATTCAATTGTTCTAGAAGTTTGTCAAGAATTGACTGTTGACATAAAATTGTAAT
 GCATAATAACAATTGAAGCACTCCTTTCTCAGTTCTCAGCTCCTCTCATTCAAGCAAATATCCATT
 TCAAGGTGCAGAACAAAGGAGTGAAGAAAATAAGAAGAAAAAATCCCCTACATTATTGGCACAGAA
 AAGTATTAGGTGTTCTTAGTGAATATTAGAAATGATCATATTCAATTGAAAGGTCAAGCAAAGACA
 GCAGAACATCAACTCATCTAGAAGTATGGAAACTAAGTTAAGGAAGTCCAGAAAGAACGCAAG
 ATATATCCTTATTTCATTCCAACAAACTACTATGATAAAATGTGAAGAAGATTCTGTTTGTGACCT
 ATAATAATTACAAACTCATGCAATGACTTGTCAAGCAAATTAAAGCAAATTATTAACATTG
 TTACTGAGGATGTCAACATATAACAATAAAATATAACACCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVC1GLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCATLTINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIIVHEKYKHPSHDYDISIAELSSPVYTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
IVSWGDECAPNKGPGVYTRVTALRDWITSKTG

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAAGTTCCCTG
 CCCCAGATGAGCCCCCGCCGTGCGTCCCCACTATCCCCAGGCGGCGTGGGGCACC GGCCAGC
 GCGACGATCGCTGCCGTTTGCCCTGGGAGTAGGATGTGGTAAAGGATGGGGCTCTCCCT
ACGGGGCTCACAATGGCCAGAGAAGATCCGTGAAGTGTCTGCCTGCTCTACGCCCTCAA
 TCTGCTCTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTCTGCTGGATGAGGGACTACCTAA
 ATAATGTTCTCACTTTAACGTGAGAAACGAGGGTAGAGGAAGCAGTCATTGACTTACTTTCT
 GTGGTTCATCCGGTATGATTGCTGTTGCTGTTCCCTATCATGTGGGATGTTAGGATATTG
 TGGAACGGTAAAAAGAAATCTGTTGCTCTTGCACTGGTACTTTGAAGGTTGCTTGTCACTTCT
 GTGAGAACTGGCTGTGGCTTTGGACATATGAACAGGAACTATGGTTCCAGTACAATGGTCA
 GATAATGGTCACTTGAAAGCCAGGATGACAATTATGGATTACTAGATATCGGTGGCTTACTCA
 TGCTTGGAAATTTCAGAGAGGTTAACGTGCTGTTAGGATAGTATATTCACTGACTGGTTGG
 AAATGACAGAGATGGACTGCCCGGAGATTCTGCTGTAGAATTCCAGGATGTTCCAAA
 CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTTGTTGGAGAAAATGTATTCTT
 TTGAGAGGAACCAAACACTGCAAGGGCTGAGGTTCTGGGAATCTCATGGGGTACACAAA
 TCCTGGCCATGATTCTCACCAATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGG
 ACAGACCAAATGATGTCCTGAAAGAATGACAACCTCTCAGCACCTGTCACTGCTCAGTAGAACT
 GTTGAACCCAAGCCTGTCAGAATCTTGAAACACATCCATGGCAAACAGCTTAAACACACT
 TTGAGATGGAGGAGTTTAAAAAGAAAATGTCAAGAAGGAAACACAAACTGTTTTATTGGACT
TGTGAATTTCAGTACATACTATGTTTCAGAAATATGTAGAAATAAAATGTTGCATAAAAA
TAACACCTAACATACATCTGTTTCAAAATGAGGATGAAAAGTTCATGTCATAAGTC
ACCACTGGACAAATAATTGATGCCCTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGGCC
TGTGATGACTTTACTGAAACACAGTTATGTTGAGGCAGCATGGTTGATTAGCATTTCGCA
TCCATGCAAACGAGTCACATATGGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATTA
CTCAGCGATCTATTCTGATGCTAACATAAAATTATCAGAAAACCTTCATATTGGTGACT
ACCTAAATGTGATTTCGTTACTAAATATTCTTACCACTAAAGAGCAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTTAAATGTTATAATTCAAGTCGAT
TTCAAGTTCTGATAATGTTAAGAATAACCAATTGAAAGAAAATTGTCTGTATAGCATCATT
ATTTTACGCTTCTGTTAAAGCTTACTATCTGCTGGGTTATATTACACATATAAC
TGTATTTAAATACTAACACTAATTGAAAATTACCACTGCTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGCTTTAGGAAGTATTAATAAGAAAATTGCACACATAACTAGTTGATT
AAGGACTTGTATGCTGTTCTCCAAATGAAGACTTTGACACTAACACTTTAAAAA
GCTATCTTGCCTCTCCAAACAAGAAGCAATAGTCTCCAAGTCATATAAATTCTACAGAAA
TAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAACACATGTGACAATTAGAGATT
CTTGTTTATTTCACTGATTAATATACTGTTGCAATTACACAGATTATTAATTTTACAA
GAGTATAGTATATTATGAAATGGAAAAGTGCATTTACTGTATTGTGATTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTGCAATAATTGTTCTAGAGAGTAA

21. 4. 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMNTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMWDWPPDSCCVREFPGCSKQAHQ
EDLSLDYQEGCGKKMYSFLRGTKQLQVLRFGLSISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACICATCCTCATCCTCTTGATAAAGCCCCTACCAAGTGCT
GATAAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTAAAAAAAAAGTGCTTGAAGAGAAGGGGACAAGGAACA
CCAGTATTAAAGAGGATTTCCAGTGTTCCTGGCAGTTGGTCAGAAGGATGCTTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGACCCCTGGCCCTAGATCCTGTTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTGGATGAGTCAGGCTCAGGCTCCCTATGTGACAACCAGTGAATGGGGAG
TGGTACCACTTCAGGGCATGGGGAGATGCCATGCCAACCTCTGCAACCCAGAAAACACTGTGGAACCCA
CCGACCTGCTGGCTCAATGGCAGCCACCCCTAGAAGGGCAGGGATTGTGCAACGCCAGGCTGTGCCAGCT
TCAATGGAACTGCTGCTCTGGAACACCACGGTGAAGTCAGGCTTGGGCTGGAGGCTACTATGTGATATCGT
CTGACCAAGCCCAGCGTCTGTTCCACGTCACTGTGGTCAATTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGACATGCCACGGACTGTCTAGGGCTGACAGGAGACATGCTTG
ATGAAAATGAATGTGAGCAAACACAGGTGGCTGAGTGAAGATCTGTGTAACCTCAAAAACCTTACCGCTGT
GAGTGTGGGGTGGCCGTGTGCTAAGAAGTGATGGAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGCTTGGATCTGAGAAAGGCTACCAAGTGTGAATGCCCCGGGCTGGTGTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTGCAAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCTGGAGCTTCCGACCAACACCTCCGTGCCAGGAGTGTCAACGGCACCCATGCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCAGTGTGGTGAATGACAAGATTGTGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCCCGGGGAGCAGGGGACTTCATCATCCGAACCAAGCAGCAAGCTGCTGATCCCAGG
ACCTGCGAGTTTCAACGCTGTACACCATTCTGAAGGATACGTTCCAAACCTCGAAACTCCCCACTGGAAAT
CATGAGGCCAAATCATGGGATCTTCCCAATTCACTCTGGAGATCTTCAAGGACAATGAGTTGAAGAGCCTTACC
GGGAAGCTCTGCCAACCTCAAGCTTGTGACTCCCTACTTGGCATTGAGCCGTGGTGCACGTGAGCGGC
TTGGAAGCTTGGTGGAGAGCTGCTTGGCCACCCCACTCCAAGATCGACGAGGCTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAAGTGGCAAGACCAAGGAAAGCTGCTTGTGACTCCGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTCAAGTTGTGGCAAGACCAAGGAAAGCTGCTTGTGACTCCCTACTTGGCATTGAGCCGTGGTCTGTGGAGTG
TTGGACGAGCGTCCCGCTGTGCCAGGGTGGCAGGGCAATGGCTGTGGGCAAGGAGGACTCAGC
CGGCTCTACAGGGCAGACGCTAACAGGGCCGAGCCATCGACTGGGAGGACTTAGTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTGCTTGGAGCTTCTCCCCCAGGGCTCAAGAACATCTGCAACAGC
TGGGTTCAAGACTTCAACTGTGAGTTCAAGACTCCAGCACCAACTCACTGATTCTGGTCCATTCACTGGGCA
CAGGTACAGCACTGCTGAACAAATGTGGCTGGGTGGGTTCACTTCTAGGGTTGAAAACACTAAACAGTACTG
CCCAGAAAAGACACTCACCCATTCCCTCATTTCTTCAACTTAAACACCTGTGATGGTCAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCCCTAGAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCCAATGACTCTTAAATATGTTAATAGTTACACCTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAAGTGTATCAATAAACAGTATATAATT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAIISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGVQRQACASFNGNCCLWNNTVEVKACPGGYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSCLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFFLTNTSCRGVSNGLHVNIIFSLKTCGTVVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRDSLFGIEPVVHV
SGLESLVESCFAPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQQQLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCACCGCAGCTTGCTCAGGGACAAGGATGCTGGGCTGAGGGACCAAGGCCGTGCCCTGCACACTCGG
 GCCTCCTCCAGCCAGTGTGACCAGGGACTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGAGGCCCT
 CCTGCTGCCCTGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACGGGAGGATCACAGAGCCAGCATGT
 TACAGGATCCTGACAGTGTACAAACCTCTGAACAGCCTCGATGTCAAACCCCTGCCAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGGAGTATCATCATTGT
 GGTTGTCTCATCAAGGTGATTCTGGATAAAACTACTTCCTCTGCAGCCTCTCCACTTCATCCCAG
 GGAAGCAGCTGTGACGGAGAGCTGGACTGCCCTGGGGAGGACGAGGAGACTGTGTCAAGAGCTTC
 CCCGAAGGGCCTGCAGTGGCAGTCGCCTCCAAGGACCGATCCACACTGCAGGTGTGGACTCGGCCAC
 AGGGAACTGGTTCTGCCTGTTGACAACACTTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
 GAGCTTCGCATGCGAACACTCAAGTGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGACTGTCTGCCIG
 TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGGAGGAGGACCTGTGGATTCTGGCCTTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCTGGACCCCCACTGGCTCAGGGCA
 GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGGGGCAGGCTCAGACAAACTGGGAG
 CTTCCCATCCCTGGCTGTGGCAAGATCATCATCATTGAATTCAACCCATGTACCCAAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTCCCACACTTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTCTTT
 GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATGGATGGGCTTACGAAGCAGAATGGAGGGAA
 GATGCTGACATACTGCTGCAGGCGTCAGTCCAGGTATTGACAGCACCGTGCAATGCAGACGATGCGT
 ACCAGGGGAAGTCACCGAGAAGATGATGTTGCAAGGCTCAGGCTCAGGCTCTGGGACACCTGCCAGGG
 GACAGTGGTGGGCCCCCTGATGTACCAATCTGACCAGTGGCATGTGGGGCATCGTAGCTGGGCTATGG
 CTGCGGGGGCCCGAGCACCCAGGAGTATACCCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
 GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCAGTGTGGAGGCCCTCTGCCCTGCCACCT
 GGGGATCCCCAAAGTCAGACACAGAGCAAGACTCCCCGGTACACCCCTCTGCCACAGCCTCAGCAT
 TTCTGGAGCAGCAAAGGCCCTCAATTCTGTAAGAGACCCCTCGCAGGCCAGAGGCGCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCCACACTTGGTGCTCCAGCATCCAGGGAGAGACACAGCCACTGAACAAGGTCT
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAAGCC
 CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCCAGGCCCTGTCCGTCTCACCCATCCCCAA
 GCCTACTAGAGCAAGAAACCAAGTTGTAATATAAAATGCCACTGCCCTACTGTGGTATGACTACCGTTACCT
 ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATTAAAGAGCTGTAAACATCTGGCAAAAAAAA
 AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVLIKVILDKYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVRVGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKI IIIIFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTQCQGDGGPLMYQS
DQWHVVGIVSWGYZGCZGPSTPGVYTKVSAYLNWIYNVWKEL

Transmembrane domain:
amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCAAAGGTGCTGTGATTATAGGTGTAAGCCACCGTGTGGCCTCTGAACAACTTTCAGCAACTAAAAAAGCCACAGGAGTTGAACTAGGATTCTGACTATGCTGTGGTGGTAGTGCTCTACTCTACCTACATTAAAATCTGTTTTGTTCTTGTAACTAGCCTTACCTCTAACACAGAGGATCTGTCAGTGCTGCTCTGGCCAAACCTGACCTCACTCTGGAAACGAGAACAGAGGTTCTACCCACACCCTCCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTGCTGGAGCAGTGCCCTCACCAACTGTCTACGTCTGGAGGACTGACTCGGGCAGTCAGGTAGCTGAGCCTTTGGTAGCTGCCCTTCAAGGTGGCTCTTGCCTGGCCGTAGAAGGGATTTGACAAGCCCGAAGATTTCATAGGCATGGCTCCACTGCCAGGCATCAGCCTTGCTGTAGTCATCACTGCCTGGGCCAGGACGGGCCGTGGACACCTGCTCAAGGCAGTGGGTGAGACATCACGCTGCCGCCACTCTAACCTTTCATGTCTGCACATCACCTGATCCATGGCTAATCTGAAACTCTGCTCCAAGGAACCCAGAGCTGAGTGAGCTGTGGCTCAGACCCAGAGGGCTGCTGAGGGTCTGTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAAACATGAGGAACGCCGGAGGAAAGCAAAGTGGCAGGGAGGAACCTGTGCCAAATTATGGGTCAAGAAAGATGGAGGTGTTGGTTATCACAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTGCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGATCCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCACAGCCAGGCAGCAGCAGCCAAATCTGCGATCACCAGCCAGGGCAGCCGCTGGGAAGGGAGCAAGCAAAGTGCACATTCTCCCTCCCTCCTCCCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACAGCAAGACATAGCTGACAGGGCTAATGGCTCAGTGTGGCCAGGGAGTCAGCAAGGCCCTGAGAGCTGATCAGAAGGGCTGCTGTGCAACACGGAAATGCCCTCAGTAAGCACAGGCTGCAAATCCCCAGGCAAAGGACTGTGTGGCTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTTCAAATGATCTCCAAGGCCCTTATACCCAGGAGACTTGATTGAAATTGAAACCCCAAATCCAACCTAAGAACCCAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACATTGGGAGGCCAGGGCGGTAGATCACCTGAGGTCAAGACACCAGCCTGGCCAACATGGTGAAACCCCTGTCTCTACTAAAAATAACAAAAACTAGCCAGGCATGGTGTGTGCTGTATCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACAGAGAATCACTCAGCCTGAGCAACACAGCAGACTCTGTCAGAAAAAATAAAAAAGAATTATGGTTATTGTAA

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGCTCTCAGTCCTCTCAAAGCAAGGAAAGACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTGTTGATCCTGGCCCTAACCTCAATTGTCCCT
GTTTGGGGAGCAAGCACTCTGGCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAACGGAAATGGCACTGATGAAACATTGAAAGTGACGACTTTAAAAACGGATAACTGG
CATCTACTCGTGGTCTTCAAAATGTTTATCAAACACTCAGATTAAAGTGATTCTGAATT
CTGAACCAGAACAGGAAATAGATGAGAATGAAAGAAATTACCAACTTCTTGAACAGTCAGTG
ATTTGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTCTTAAACATTCCAAATTCTGGA
GATTGTGATAACGTGACCATGATTGGATCAATCCACTCTAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGAGGGAGAAGATCTCACTTCCCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTGAAAGTAGAGAACGCCCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCAGCTGGATGAGAGAG
GTTATTGTTGATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTGTGAACCTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATGCCTTG
TAACTGGTGGTGGCCCGCATGCTGGGGAGGGTCTAAATAGGAGGTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGAGGTAATTCTCTTCAATGTTCTAATAAAACTCTACA
TTATCACCAAAAAAAAAAAAAAA

HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEETTTFFEQSVIWVPAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE
LQDEEEEGEDLHFPAKEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTACACCTTCGGCACAGGAGGGCGGCAGCTCTCGCAGGCCA
 GGGCGGGCGGCCAGGATCATGTCACCACATGCCAAGTGGTGGCTTCCCTGTCCATCTGGGGCT
 GGCGGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
 CCGTGTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGAGAGTTCAAGGCTTACCGAATGCAGGCC
 TATTTCACCATCTGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
 GGGTGCCTGGCATTGGCTCTGGTATCCATCTTGCCTGAAATGCATCCGATTGGCAGCATGGAGGACTCTG
 CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTCAATTGTCAGGTCTTGTGCAATTGCTGGAGTG
 TCTGTGTTGCAACATGCTGGTACTAACCTCTGGATGTCACAGCTAACATGTACACCGGCATGGTGG
 GATGGTGCAGACTGTTAGACACAGGTACACATTGGTGCAGGCTCTGGTGGCTGGGCTGGGCGCTGGAGGCC
 TCACACTAATTGGGGTGTGATGATGTCATGCCCTGCCGGGCTGGCACAGAGAAAGAAACCAACTACAAA
 GCCGTTCTTATGCTCAGGCCACAGTGTGCCTACAAGCCTGGAGGCTTCAGGCCAGCAGTGGCTT
 TGGGTCACACAAAAACAAGAAGATATACTGATGGAGGTGCCGCACAGAGGAGGAGGTACAATCTTATC
 CTTCCAAGCAGCAGTGTGTAATGCTCTAACGACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
 CCCAAAAAACAGGAGATCCCCTAGATTTCTCTGCTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
 CGATTTCATCTTGGAGAGGCCAAATGGCTTAGCCTCAGTCTGTCTAAATATTCCACCATAAAACA
 GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTTCATCCTTAAATTTTTAAATATAACT
 TTCTACTCTGATGAGAGAATGTGGTTTAATCTCTCCTACATTGATGATTAGACAGACTCCCCCTC
 TTCCCTCTAGTCATAAACCCATTGATGATCTATTCCAGCTTATCCCAGAAAACCTTCTACTGCTG
 GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGCTCCACCTGGGCTAGTAACTGCTGTT
 ACACCTACTGAAGAAGAAGCAATAAGAGAAAAGATATTGTAATCTCTCCAGCCCATGATCTCGGTT
 ACAGTGTGATCTAAAGTACCAACCCATTGATGATCTATTCCAGCTTGTGAGGCAACCAACCTTCTACTGCTG
 TTGACATCTCTTATTACAGCAACACCATTCTAGAGGTTCTGAGCTCTCCACTGGAGTCTCTTCTGT
 CGGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTAAATTAAAGCTTAAATATAGTTAA
 AATAAATAATGTTTAGTAAATGATACTATCTCTGAAATGCCCTACCCCTACATGTGGATAGAAG
 GAAATGAAAAAATAATTGCTTGACATTGCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAATTCC
 ATGAAAAGCTCACACCTGTAATCCTAGCATTGGAGGCTGAGGAGGAAGGATCAGTGGAGCCAGAAGT
 TCGAGACTAGCTGGGCAACATGGAGAAGCCCTGTCTCACAAAATACAGAGAGAAAAATCAGCCAGTC
 TGGTGGCATACACCTGTAAGTCCCAGCATTCCGGGAGGCTGAGGTGGAGGATCAGTGGAGCCAGGGAGGT
 TGGGGCTGCAGTGAGCCATGATCACACCAGTGCAGCCAGGTGACATAGCAGGATCCTGTCTAAAAA
 AATAAAAAATAATGAAACACAGCAAGTCTAGGAAGTAGGTTAAACTAATTCTTAA

MSSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRPYFTIILGPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHHSVAYKPGGFKA
STGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRPYFTIILGPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIACRGLAPEETNYKAVSYHASGHHSVAYKPGGFKA
STGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
 AGCTGGCTAAACATCCCAGAGGATAATGGCAACCATGCCCTAGAAATCGTGGCTGTTCTTG
 GTGGTGTGGAATGGTGGCACAGTGGCTGTCACTGTCAGTGGCTAGTGGAGAGTGTGCGGCTTCATT
 GAAAACAACATCGGGTTTGAAAACCTCTGGAAAGGACTGGAATTGCGTGAGGCAGGCTAA
 CATCAGGATGCAGTGCAAAACTATGATCCCTGCTGGCTTTCTCCGGACCTACAGGCAGGCCAGAG
 GACTGATGTGTGCTGCTTCCGTATGTCCTCTGGATCATGATGCCATCCTGGCATGAAATGC
 ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTCAT
 CATCACGGGCATGGTGGCTCATCCCTGAGCTGGGTCATGCCATCATCAGAGATTCTATA
 ACTCAATAGTGAATGTTGCCAAAACGTGAGCTGGAGAAAGCTCTACTTAGGATGGACCACGGCA
 CTGGTGTGATTGTTGGAGGAGCTGTCTGCTGCGTTTTGTGCAACGAAAAGAGCAGTAGCTA
 CAGATACTCGATACCTCCCATCGCACAACCCAAAAAGTTATCACACCGGAAAGTCACCGAGCG
 TCTACTCCAGAAGTCAGTATGTTAGTTGTTATGTTTTAACTTACTATAAGCCATGCAAATG
 ACAAAAATCTATATTACTTCTAAAATGGACCCAAAGAAACTTGATTACTGTTACTGCCT
 AATCTTAATTACAGGAACTGTGCATCAGCTATTGATTCTATAAGCTATTGAGCAGAATGAGATA
 TTAAACCCATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTTCAAGCATTCTA
 CTCTTTTATCATTACTCAAAATGACATTGCTAAAGACTGCATTATTACTACTGTAATTCTCC
 ACGACATAGCATTATGTACATAGATGAGTGTAACTTATCTCACATAGAGACATGCTTATGGT
 TTTTTAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTAGGGAA
 ATCATGGATAGGGTGAAGAAGGTTACTATTGTTAAAACAGCTTAGGGATTAATGCTCTCCA
 TTTATAATGAAGATTAAAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
 GCTGTTTTAGCCTAGGAGTTAGAAATCTAACTTCTTATCCTCTCCAGAGGCTTTTT
 CTTGTGTATTAAATTACATTAAAACGCAGATATTGTCAGGGCTTGCATTCAAACTGCTT
 TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTAGGAAAGTG
 AAAATTTTAAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAATAAAAGAGCAGAAAATA
 TGCTTGTTCTATTGCTTACCAAAAAACAAACAAAAAAGTGTGCTTGTGAGAACCTCACCT
 GCTCCTATGTGGGTACCTGAGTCAAATTGTCATTGTTCTGTGAAAATAAATTCTCTTGTA
 CCATTCTGTTAGTTACTAAAATCTGAAATACTGTATTCTGTTATTCAAATTGATGAA
 ACTGACAATCAAATTGAAAGTTGTGTCGACGTCTGCTAGCTAAATGAATGTGTTCTATTGCTT
 TATAACATTATTAATAAAATTGACATTCTAATT

© 2007 International Society for Computational Biology. All rights reserved.

FIGURE 120

MATHALEIAGLFLLGGVGMVGTVAVTVMQPQRVSIFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSSLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRGTGNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTTLVLIVGGAFCGVCCNEKSSSY
RYSIPSHRTTQKSHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCCGCGGGTCAAAGGCATTGATGCAGCCTGCAGCCCTCGGAGGCCGGAG
CCAGACGCTGACCACGTTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCGAGCC
GGGAGCCATGCGACCCCAGGGCCCGCCCTCCCCGAGCGGCTCCGCGCCCTGCTGCTCC
TGCTGCTGCAGCTGCCGCCGTCAGCGCCTCTGAGATCCCCAAGGGAAAGCAAAAGGCCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGTTACAAGGGCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCAATGTTATTCCGGTACACCTGGGATCCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCAGGGAAAGCTTGAGGAGTCCTGGACACCCAACTAC
AAGCAGTGTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGGAGTGTCATT
TACAAAGATGCGTTCAAATAGTGCTAAAGAGTTGGTCACTGGCTACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTCACATTCAATGGAGCTGAATGTTCAAGGACCTTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAAATTATCG
CACTCTTCTGAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTCTACTGGATGAAATTCACTGGCTTC
ATTATTGAAGAACTACCAAATAAATGCTTAATTTCATTGCTACCTTTTTATTATGCC
TTGGAATGGTCACTAAATGACATTAAAGTTATGTATACATCTGAATGAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTCACACTGTTAAATCTAGCATTATTCAATT
CTTCATCAAAGTGGTTCAATATTGTTAGTTGTTAGAATACTTCATAGTCACATT
CTCTCAACCTATAATTGGAATATTGTTGTTAGTTCTTAGTATAGCATTAA
AAAAAAATATAAAAGCTACCAATCTTGACAATTGTAATGTTAAGAATTTTTTATCTGT
TAAATAAAATTATTCCAACA

MRPQGPAA SPQR LRGLL LQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR

DGSPGANV I PGT PGI PGRDGF KGEK GECL RESFEE SWTP NYKQC SWSSL NYGID LGKIA ECT FTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECS GPLPIEAI IYLDQGSPE MNSTINI HRTS
SVEGLCEGIGAGL VDVAI WVGTCSDYPKG DAST GWN SVSRIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCTGGAAAGAATACATCATGTTTCGATAAGAAGAAATTGAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTGGATGTTGCTGCGAATGCGGTGTGGGATTATTGTTCTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTCAAAATCGGCCATCTCCAAGGGTCCAATT
TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGCTGTATGCACTG
GCCCTAAGCAAAGCAAAGACCTAAGGACGACCTTGAAACAATACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
CTTTCTCTGCCAACGAGATGCCCTAACGGCTGTAGGTGTGAAGGCAAATGGTATATTGTGA
ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTGTCCTCGCT
ATAACAGCCTCAAAACTTAAGTATAATCAATTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATAACCATACTAGCAATATTGACGAAATGCTTTAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAATCTCTATTTCCTAACAAATACCTCAGACCTGTGACAAATT
TACGGAACCTGGATCTGCTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGCTTAACTCCCTGAGAACCATCCCTGTGCAATATTCCA
AGACTGCCGAACTCTGAACTTTGGACCTGGGATATAACCGGATCCGAAAGTTAGCCAGGAATG
TCTTGCTGGCATGATCAGACTCAAAGAACCTCACCTGGAGCACATCAATTTCACGCTCAAC
CTGGCCCTTTTCCAAGGTGGTCAAGCTTACTGGGATCTGGGCTGAGGAAATTTACAGTGT
CATAGGACAGACCATGTCCTGGACCTGGCTAACAAAGGCTGATTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCTGAACTGCAAGCCCTAACCTGGAT
TCCAACAAGCTCACATTATTGGCAAGAGATTGGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAAATATGGGAAATGCAAGGAGAACATTGGCTCCCTGTAAACTGGCTGAAAGTT
TTAAAGGTCTAAGGGAGAAACAATTATCTGTGCCAGTCCAAAGAGCTGCAAGGAGTAAATGT
ATCGATGCAGTGAAGAAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTGATCTGGCCAG
GGCTCTCCAAAGCCGACGTTAACGGCCAAGCTCCCAGGCCAACGATGAGAGCAAACCCCCCT
TGCCCCCGACGGTGGGAGCACAGAGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
TTCCATAAAATCATCGGGGAGCAGCTGGCTTTCTGTCCCTGCTGTATCCTGCTGGTTAT
CTACGTGTCAAGGAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCCGCTCCCTCATGC
GGCACAGGAAAAGAAAAGACAGTCCCTAACAGCAATGACTCCAGCAGCCAGGAATTATGTA
GATTATAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCTGCAAC
TAACAAATCGGGTCCAGGGAGTGTGAGGTTGAACCATTGTGATAAAAAGAGCTTAAAGCT
GGGAAATAAGTGGTGTCTTATTGAACTCTGGTGAATCAAGGGAACCGCATGCCCTCCCC
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTCCTGTCCGTTAGTGCATTCAATAACT
GGTCATTTCCTCTCATACATAATCAACCCATTGAAATTAAACCAACATCAATGTGAAGCTT
GAACCTGGTTAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTAAATGTCGCATT
GTTTAAGATAAAACTCTTCATAGGAAAAAAAAAA

FIGURE 124

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGCTTGCCTACTGCTGAATGTCCCGCCGGAGGAGGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGCGGGCTACCGTGGCCGAGCT
AGCAACCTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAACAGCTTGCCTC
GGTTGGGAGACGGTGCAAGAGAACTGCCCCATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGCACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTATTGGCCAGTT
TTAGCCAATCCAACTGACCTAGTGAAGGTTCAGATGCAAATGGAAGGAAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCTGTTGACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGCGCTGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTAA
ACCACTTATGATACAAGTGAACACTACTTGGTATTGAATACACCACCTGGAGGACAATATCATGAC
TCACGGTTTATCAAGTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGGACTTTGTATAAAATCATCG
ACTGACTGCTTGATTCAAGGCTGGTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTAC
ATCTTGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLEVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTYDTVKHYLVNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRI MNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGGCCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTGGGCTAGTCAGCGAGGCCAGGGCGCTGGCGTGGTGACGGGGCCTGGGGGGCATGGCGCGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGGCCCGACTGTGGGAAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTCAGGCATACTGGGACTTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTGGCCGGCTGACACCTGCTCTCAGGCAGCACCAAGTGGTTGGAA
GGACATGTTCAATGTAACGTGGCAGTGGCACATCATTAAACATCAATAGCATGTCAGTGGCCACCGAGTGTAA
AGGAGCGGAATGTGGACGATGGGACATCAGTGGCACATCCGAGCCACGAGCTGGCCACCGAGTGTAA
CCCCGTCTGTGACCCACTTCTATAGGCCACCAAGTATGCCGTCAGTGGCCACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACGAGCTGGCCACCTATGAGCAA
AGACACAATTGCCCTCAAACCTCCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTAAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCTCAGCACCCCCGACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACC TAGTGA
CTGGCTCTGGCTTGTGACCCACTTCTGTCACACCCCGACCAGGGCTAGAAAATTGTTGAGATTTATA
TCATCTTGTCAAATTGCTTCAGTTGAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTGTGCCCTGGGACTTGGCCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGGCCAAATCCCCATCTTCTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCCTGGCCACTGCACCCCTCCCCCTATCTATCTCCTCTCGGCTCCCC
AGCCCAGTCTGGCTTGTCCCTCTGGGTGATCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTGATGTTAAAAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRILALVTGASGGIGAAVARALVQQGLKVVGCARTVGНИEELAAECKSAGYPGTЛИ
PYRCDSLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRLPLSVTHFYSATKYAVTALTEGLRQEЛREAQTHIRATC
ISPQVVETQFAFKLHDKDPEKAATYEQMCKLKPEDVAEAVIYVLSTPAHQIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACCTCTACATGGGCTCCGTGCTGGTCTTCCTCAGCCTCTGCCGGTGGCCTACACCAT
CATGTCCTCCCACCCCTCCTTGACTGCGGGCGTCAGGTGCAGAGTCTCAGTTGCCGGAGC
ACCTCCCCTCCGAGGCAGTCTGCTCAGAGGGCCTGGCCCAGAATTCCAGTTGGTTCATGC
CAGCCTGAAAAGGCCATGAACTTGGGTGAATCACCGATGCCATTAAAGAGGGTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTCTCGCCTGTCAGTTCACTGAGCCACAGCACGTGG
CCGTTGAGTGCTTGAATGAGAACTGAGAAAATTAAATTCTCATGTATTTCATTTCTATTAA
TTAATTAACTGATAGTTGACATATTGGGGTACATGTGATATTGGATACTGTATACAA
TATATAATGATCAAATCAGGGTACTGGGATATCCATCACATCAAACATTATTTTATTCTTT
TTAGACAGAGTCTCACTCTGCACCCAGGCTGGAGTGCAGTGGGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAAAGTAGCTGGACTACAGGCAT
GCACCACAATGCCAACTAATTGTATTTTAGAGACGGGTTTGCCATGTTGCCAGG
CTGGCCTGAACCTCTGGCCTAAACAATCCACTTGCCCTGGCCTCCAAAGTGTATGATTACA
GGCGTGAGGCCACCGTGCCTGGCTAAACATTATCTTCTTGTGTTGGAACTTGAAATTAT
ACAATGAATTATTGTTAACCTGCTCATCTCCCTGCTGTGCTATGAAACACTGGGACTTCTCCCTCT
ATCTAACTGTATATTGTACCAAGTTAACCAACCGTACTTCATCCCCACTCCCTCTATCCTCCCC
AACCTCTGATCACCTCATTCTACTCTCACCTCCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCTTCTGTGCCTGGCTATTCACTAACATAATGACTCCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

MGLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSSRGSSLRGPRPRIPVLVSCQPV

FIGURE 130

MGLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSSRGSSLRGPRPRIPVLVSCQPV
KGHGTLGESPMFKRVFCQDGTVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAAACGGAAGCTACCTGTATAAAGACCTCAACACTGCTGACCTGATCAGCCGAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGCTGTTCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCTG
TCCATCTGTGTCGCTGCGATGCCGGTTCTATTACTGTAATGATCGCTTCTGACATCCATTCAAACAG
GAATACCAAGAGGATGCTACAACCTCTCACCTCAGAACAAACAAATAATGCTGGGATTCCCTCAGAT
TTGAAAAACTGCTGAAAGTAGAAGAACATACCTATACCACAAACAGTTAGATGAATTCCCTACCAACCT
CCCCAAAGTATGTAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGGAGAACATTACATTAGATGACAACCTCTGCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTCTGTCGGTAAACACCTTAGCACAAATTCCCTGGGGTT
GCCCAAGGACTATAGAACGACTACGCCGGATGATAATCGATATCCACTATTCATCACCACCTCTTCAA
GTCTCACTAGCTAAACGCCCTGGTCTAGATGAAACCTGTTGAAACAATCATGGTTAGGTGACAAAGT
TTCTTCACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGAATTCCCTGACTGCTGCACCAAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTTCAAGATAACCCATCAATGGGGCCCCAAATGCTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGCCAATAAACCTAACGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGTATTCTCGCAACAACTCCCTGGTATTGGCTGCAAGATGAAATG
GGTACGTGACTGGTACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGGGGATGGCTATTAAGGATCTCAATGCAGAACACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCAACTGCAATACCCAAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCAGTGAC
CAAACAGCCAGATATTAAGAACCCCCAGCTCAACTAACGATCAACAAACCCAGGGAGTCCTCAAGAAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTATCTTGGAAACTTGTCTACCTATG
ACTGCTTGGAGACTCAGCTGGCTTAAACTGGGCATAGCCCCCATTGGATCTATAACAGAAACAAATTGT
AACAGGGGAACGCGAGTGGACTCTGGTCACAGCCCTGGAGCCTGATTCCACCTATAAAGTATGCATGGTC
CCATGGAAACCCAGCAACCTCTACCTATTGATGAAACTCTGTTGTATTGAGACTGAAACTGCAACCCCTT
CGAATGTACAACCTACAACCAACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCATTAC
TTTGGCTGCCATCATGGTGGGCTGTGGCCCTGGTACCCATTGCCCTTGTGTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTCTCAAGGAACACTGTCATATGCAAAGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACTCTATCCTGGAAATCAGGGAAACTTCTTCACTGTTACCAAT
AAGCAATGAACCCATCTGAAGGAGGAGTTGTAATACACACCAATTCTCTCTAATGGAATGAATCTG
ACAAAAACAAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCTGATGCTGAAGGACTCACAGCAGACTTGTGT'TTGGGTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAFGIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSIDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSDLスキPYL
EELHDDNSVSAVSIEEGAFRDSNYLRLLLFSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSSL
QGLTLSKRLVLDGNLNNHGLGDVKFFNLVNLTESLVRNSLTAAPVNLPGTNLRKLYLQDNHN
RVPPNAFSYLRQLYRLDMSNNNLSNLPGIFDDLDNITQLILRNNPWCGCKMKWVDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDGSIVSTIQTTAIPNTVPAQGQWPAPVTKQPD
IKNPKLTKDQQTTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVMVPMETSNLYLFDETPVCIETETAPLRMYNPTTLNREQEKE
PYKPNPLAIIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNS
ILEIRETSFOMLPISNEPISKEEFVIHTIFFPONGMLYKNNHSESSSNRSYRDGGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site

amino acids 567-570

Leucine zipper pattern

amino acids 159-180

Phospholipase A3 aspartic acid active site

amino acids 34:41

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCAGAGTCCTTGCCAGGCCACCCAGGCTTGGCA
 GCCCTGCCGGGCCACTGTCTTCCATGTCTGCCAGGGGAGGTGGAAGGAGGTGGGAGGAGGGCG
 TGCAGAGGCAGTCTGGCTGGCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGACAGAG
 GCCGGCATGGCCAGCCTGGGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
 CCTCCTCACTGCCCTGGGCTGGACACTGCTGAAAGTAAAGCCACATTGACAGACCTGATCCTGTCT
 GCGCTGGAGAGAGCCACCGTCTTCTAGAACAGAGGCTGCCGAAATCAACCTGGATGGCATGGT
 GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGAGAAGTGGGCCAGGAGCCCCTGC
 TGCAGCCGCTGAGCCTGCGCGTGGGATGCTGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
 CTCCACTACCTCAAGCTGAGTGTACCTAACGACTGATGCCCTGGTGTACCCCACGTTGGGC
 GTTTGGAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTGGTGTACCCCACGTTGGGC
 CCCAGGACTCATTCTCAGAGGAGAGAAGTGTACGCTGCCGCTGGTGTACGCTGGGAAACGGGACG
 GACAGCAGCGAGCCCTGCGCCTCTCACGACTCTGAGGAGCCATGACCAAGCCGGCTGCTC
 AGGCTACTGCCGTGCCCACCAACTGCTCTTCTGGGAGAATGAGGGATGCACACAGG
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTGCGCCAACATGATGGACTTGAACCGC
 AGAGCTGAGGCCATCGGATACGCCCTACCCGGACATCTCATGGAAAACATCATGTTCTG
 TGGAATGGGGGCTTCTCCGACTTCTACAAGCTCCGGCTGGAGGCCATTCTCAGCTGGCAGA
 AACAGCAGGAAGGATGCTCAGGGAGCTGAGTGAAGAGGCGAGAAAACAATTCCAGATTCTCGCTCTG
 TATCAGCAGCATTTTCGAGGAGAGTGAAGAGGCGAGAAAACAATTCCAGATTCTCGCTCTG
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGGCTCTGGGTTCAAGC
 AATTCTCTGCCCTACCTCCCGAGTAGCTGGACTACAGGAGCGTGCACCCATACCTGGCTAAT
 TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGCTCGAACCTCTG
 CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGATTTAGGTGTGAGCCACCGTGTCTG
 GCTGAAAAGCATTCAAAGAGACTGTGTTGAATAAGGGCCAAGGTTCTGCCACCCAGCACTC
 ATGGGGCTCTCCCTAGATGGCTGCTCTCCACACAGCCACAGCAGTGGCAGCCCTGG
 GTGGCTTCTATACATCTGGCAGAATACCCCCCAGCAAACAGAGAGGCCACCCATCCACACCG
 CCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCGGAGGAACAGACCC
 TTAGTCCTCATCCCTAGATCTGGAGGGCACGGATCACATCTGGGAAGAAGGCATCTGGAGG
 ATAAGCAAAGCCACCCGACACCCAACTTGGAAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
 GGCCGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAAAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERVTSSEQRPAMASLGLLLLLTLPPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQLRPEINLDGMGVVRVLEELQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYPLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTCTTCCGCTGCTGCTG
CTGCTGCTATCGGGGGATGTCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGCCGAGTGCTGGTAGACGGAGAACGACGTCGGTTCTTAAG
ACAGATGGGAGTTTGTGGTCATGATATACTTCTGGATCTTATGAGTTGAAAGTTGTATCTCC
AGCTTACAGATTGATCCCCTCGAGTGGATATCCTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCAAATGAAATCTCAGGT
CCACCTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTCTCCTTATTGATATTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTCTCTCAAATCATCTGGCAAATCTAGCAGCGGAGCAG
TAAAACAGGAAAAGTGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTTGGTACACGAGAAAACCCAGCTTCATCTTGCTGTAT
GAGGTCAATATTGATGTCACTGAATTAAATTACAGTGTCTATAGAAAATGCCATTAATAAAATTAT
ATGAACTACTATACATTATGTTATTAATTAAAACATCTTAATCCAGAAATCAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFVPLLLLLLSDVQSSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKGSGPPSYFIKRESWGWTDFLMNPMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKGSSKGSSKTGSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTGTGAGATTGATTTCTCCCAGTTCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCAGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
GTTTGACATTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCTGCTGTTCCAGGC
CTTACCTGCTGGGACTAACGGCGGAGGCCAGGATGGGACAGAATAAAGGAGGCCACGACCTGTGC
CACCAAUCTCGCACTCAGACTCTGAUACTCAGACCTGAAATCTCTTCACGGGAGGCTTGGCAGT
TTTCTTAUACTCCTGTGGCTCCAGATTCAAGGCTAAGATGAAAGCCTTAGTCTGCCTTCAGC
CTTCTCTGCTGCGTTTATCTCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGATGCCAACAAACCTTCAGGAAUACGAAATGGATTTCTGAGATACGGGCAGTG
TGCAAGCCAAGATGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAUACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTGTGAAGGC
TTGGGGGAUACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGCCAA
ACCACCATCTCTTACTGACTAGTCTTGCTGGCACAGTGTATCTTATTGATTGCTTCTG
CTTCCTTGCAATTGCTTATGCAATCCCAATCTTAATTGAGACCATACTGTATAAGATTT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTGCTATTAA
ATGTATTATTTTTACTGGACATGAAACTTIAAAAAAATTCAAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTTATACAGTAAAAAAAAACCTGTAAATTCTAGAAGAGTG
CTAGGGGGGTATTCAATTGCTTCAACTAAGGACATATTACTCATGCTGATGCTGTGAGAT
ATTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATATTGTGATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTCTGCATACCAAAAAAA
AAAAAAAAAA

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTATNSHSDSELRPEIF

SSREAWQFFLLLWSPDFRPKMKASSLAFLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLD RVFKNYQTPDHYTLRKIS
SLANSFLT IKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTCCGCATCCAGCC
 TAGCGTGTCCACGATGCGGTGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
 TAGGGATCGCGGTCTTCCTCGGGATTCTTCCCGCTCCGTCGTTCTCTGCCAGAGCGGAA
 CACGGAGCGGAGCCCCCAGCGCCGAACCTCGGCTGGAGCCAGTCTAACTGGACCACGCTGCC
 ACCACCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTGTTG
 GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTTGAGAAAAGGAGCATCTCACAGT
 TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGGAG
 CCTTCCCTGGCTTGTGCGACGTCACTCGAACCTCAATTCTCTGCACTGCTGGAAAGACAGTGTGA
 TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAAATTA
 TTCCCAAAGATTGTTGAAATGATGAAACCTCATTTCTGTCAGATTACACAGAGGT
 GGATAATAATGTCACGAGGCATTGGATAAAAGTATTAAAAAGAGGAGATTGGGACATATTAAATCC
 TCCACTACCTGGGCTGGACCACATTGCCACATTCAAGGCCAACAGCCCCCTGATGGGAG
 AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCACTGCAAGGGAGAGA
 GACGCCCTTACCCAATTGCTGGTTCTTGTTGACCATGGCATGTCGAAACAGGAAGTCACG
 GGGCCTCCACCGAGGGTGAATAACACCTCTGATTTAATCAGTTCTGCGTTGAAAGGAAA
 CCCGGTGAATCCGACATCCAAGCAGTCCAAATGACGGATGTGGCTGCGACACTGGCGATAGC
 ACTTGGCTTACCGATTCACCAAAGACAGTGTAGGGAGCCTCTATCCCAGTTGAGGAAAGGAGAC
 CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTGCAGCTTAGTAAACTGTCAGAG
 ATGTCGCGTCAATGAAAAAGATCTGGGTTTGAGCAGTTAAATGTCAGAAAGATTGCA
 GAACTGGATCAGACTGTACTTGGAGAAAAGCATTAGCAGTCTATTCAACCTGGGCTCCAAGG
 TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTGTCCTGAGTGCACAAAGTGGCCAG
 TTCTCACCCCTGCTCCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAAGTCCA
 CTGTCATCTCTGGTTTCTGCTCTTTATGGTGTGATCTGGTTCTTCGGCGTTCACGT
 CATTGTGTGCACCTCAGCTGAAAGTTCGACTCTGTGGCTCTCGTGGCTGGCGCAGGCT
 GCCTTCGTTAACGACTCTGGTGAACACCTGGTGTGCGCAAGTGTGGCAGTGCCTGGAC
 AGGGGGCCTAGGGAGGACAGTGGAGCAGCCTTATCCCAGGCCCTGGGTGTCCCACAGGTG
 TTCACATCTGTGCTGTCAGGTCAAGATGCTCAGTTCTGGAAAGCTAGGTTCTGCACTGTAC
 CAAGGTGATTGAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
 TCGGACAGCCTCCAGCAGGGTGTGGAGCTGCAGCTGAGGGAAAGAGACAATCGGCCCTGGA
 CACTCAGGAGGGTCAAAGGAGACTTGGTCGACCAACTCATCTGCCACCCAGAATGCATCCT
 GCCTCATCAGGTCAGATTCTTCCAAGGCCAGTTCTGTTGAAATTCTTAGTCCTGGCC
 TCGGACACCTCATTGTTAGCTGGGAGTGGTGGTGGAGCAGTGAAGAAGAGGGCGATGGTCAC
 ACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCAGTGCAGCAGGACTGTTGGGCCCC
 ACCCCAAACCCCTGCACAGGCCCTCATCCCCCTCTGGCTTGAGCCGTGAGGGCCCTGCTGAGTGT
 CTGACCGAGACACTCACAGCTTGTCACTAGGGCACAGGCTTCCTCGGAGGCCAGGATGATCTGTG
 CCACGCTTGCACCTCGGGCCATCTGGCTCATGCTCTCTGCTATTGAAATTGAGAAAAAAA
 CTGCACACAGTATGAGTTACCAAAAGAATAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVATEVLGIAVFLRGFPAPVRSSARAEHGAEPAPAEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDHSVIRQAKAAGKRIVFYGDETWKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCAGCAGGCAAGCCTTCCAGTTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGTGGCACTGATGTACTGCTGCTGGAGTACA
ACTCCCTATAGAAAACA
CTGCCAGCACCTTAAGACC
ACTCACACCTCAGAGTGAAGAA
CTTAAACCCGAA
AGATTCA
GCAGTCAGATA
AAACTACATACGCCAGAGATCT
CTTGCA
TAGCCTCATC
CTTGAGCT
CAGCCTCTGCGGAGAA
AGGAAGTCCGATT
CTCCTGGGGTCT
CTAAAGGGAG
TTTGTCT
TAC
TGTGACAAGGATA
AAGGACAAAGT
CATCCATCC
CTCAGCT
GAAGAAGGAG
AAACTGAT
GAAGCT
GGCTGCC
AAAAGGA
ATCAGCAC
GCCGG
CC
CTCAT
CTT
TATAGGG
CTCAGG
TGG
CTC
CTGG
GA
ACATG
CTGGAG
TCGG
GGCT
CAC
CCGG
ATGG
TCT
CAT
CTG
CAC
CT
CTG
CA
ATT
GT
AAT
GAG
CT
GTT
GGGG
TGAC
AGATA
AA
TTTGAG
AACAG
GAA
ACAC
ATT
G
AAT
TT
TC
ATT
CA
ACC
AG
TT
G
CAA
AGCT
GAA
AT
GAG
CCCC
CAG
TGAG
GT
CAG
CG
ATT
AGG
AA
ACT
GCC
CC
ATT
GA
AC
GC
CT
CT
CG
CTA
ATT
GA
ACT
ATT
GT
ATA
AAA
AC
CCA
AA
ACT
TG
CT
ACT

FIGURE 142

MLLLLEYNFPIENNQHLKTTHTFRVKNLNPKKFSIHDDQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGACTATAAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCCTCCCTTAA
TCCAGGATCCTGCTCTTCTGTCTGTAGGAGTGCTGTGCCCAGTGTGGGGTGGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGGCTGCCGTGG
AAGGGAGGTCTGCTCTGGCGCTGCTGCTGCTTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTGTTACCTCCCTGCCA
CACTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCCTACAGACTGGGG
CACATCTGAATTCCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGAGAACTAGGTGTGGGAAATTGAAGACGACATTGACAACGTGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTACCTGCTTCTCACCATCAGCACCAAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGAGGGATTCCACTTGATGAAACCCACTCACAGGCTTGCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCTTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAA
AAA

MLGLPWKGGLSWALLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

FIGURE 144

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGGCACTGTCTGCATGCTGCTTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAACATCAAGTGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCAAGCTCGACATT
GACTCGGAGCCAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCTGCATCAATGCCA
CCCAGGCCGAACCAGGGGAGTTCCAGAACGCCAGAACAGCTCCACCAAGGTGCTCTGG
CGGCTGGTCCAGGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGAGAGGGGCGCAGGACT
TCGGGTCAACATGCACCAAGCCAGTGCTCTGCTCTGGCTTGATCTGGCTCATGGTGAAAAT
AAAGCTTGCAGGAGGCTGGCAGTACAGAGCGCAGCAGCAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCCACGTATGCCCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGACTGTGGCTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGATAATGTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTAAGTACTTCCAACAACACTTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAAGCTAGAAAGTGGCAGGCCAGGATTCAACCCCTGGCTGTCAACCCCAGGTTCTGCTCT
CTCTAATGAAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWRKALPSTAQITEAQVAENRPGAFIKQGRK

LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNLHQQ
VLWRLVQELCSLKHFCEFWLERGAGLRTMHQPVLLCLALIWLKV

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCAAGTCAGAGTCATTT
CCTGATGATTTAGACTCAAAGAAAACTCATGTTCAGAAGCTCTTCTCTGGCCTCCTCT
CTGCTTCTTCCCTTTCTTCTTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGGTAGGTCACTGAAGTCTAGTTTTATTTTGAAATT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

© 2000 University of Michigan Health System

FIGURE 148

MFRSSLLFWPPLCLLSLFLILLISIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGGTACAGGAACCTCAGCACCCACAGGGGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCCGGGCCCCAACCCCTGTTATCCCTGACCGTCGAGTGTCAAGAGATCTGCAGCCGCCAGTCC
CGGGCCCTCTCCGCCAACCCCACCCCTCTGGCTTCTCTGGTTTACTCTCCTTTCTTCATA
ACAAAAGCTACAGCTCCAGGAGCCCAGGCCGGCTGTGACCCAAAGCCAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCACTTGGATTCTGGTGTAGTGCTCCGATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTACATAATAGAGAATTAGTGAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAACATATCCTCCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGAAAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAGAGTGGATTGATCATAAATTCAAGATGATCCAGATGGCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCATAAAATCGTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTCGGCCATTACAGAAAGC
CAAGCACATACACTGAAAGATGAAGTAGCAGAGGTTTACAAAATTATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTCCAAATTCTATGCGCTACTGAAAGTATTGATTCAAGAAAAGAGCAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTAAGGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTCTACCTTGAAAGCAGATGATGGTGAATGATTGCTTCAAGACAAAAAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAACAGATGGAAAAGGAATATGGAAGCTGTAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGCCTATTGGAAGC
CATCAGAAAAAATTGAATGGTGAAGAAACATGACAAAAGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGATTAGCAGCCTGTAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAATTTTGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTACAAGTGGTAAAACATAGCTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAA
AAAAAAAAAAAAAAA

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCCTGAGTGGGCGTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCACTAAAGGTGAAGAGATCAGC
GTGGTCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTATCCTGGGTGTCAGGGTGGAG
CCAGTGCTGTATGTGGGGTGGGGCAGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTGGTGCCAAGGAATCCAAGAGCTCACCTCTACCGGGGGACATGGGCTCACC
TCCAGCTCGAGTCGGCTGCCAACCCGGCTGGTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCACTACCCAGCTCCCGAGAATGGTGGCTGGAATGCCCATCACAGACTCTACTTCC
AGCAGTGTACT**TAGGGCAACGTGCCCCCAGAACACTCCCTGGGAGGCCAGCTGGGTGAGGGT**
GAGTGGAGGAGACCCATGGGGACAATCACTCTCTGCTCAGGACCCCCACGTCTGACTTAG
TGGCACCTGACCACTTTGTCTTCTGGTCCAGTTGGATAATTCTGAGATTGGAGCTCAGT
CCACGGTCCTCCCCACTGGATGGTCTACTGCTGTGGAACCTGTAAAAACCATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGGGGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCGCAGGCCAACCCATCCCCAGTTGAGCCTTATA
GGTCAGTAGCTCTCACATGAAGTCTGCACTCACCACTGTGAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCCCTTAATCCTGCCACTGTCTA
TGCTACCTTCTATCTCCCTCATCATCTTGTTGTGGCATGAGGGAGGTGGTGTAGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTGGGTATGCTGATCCTCTTTAAAAACCAA
GATAACATCAAATCCAGATGCTGGTCTCTATTCCATGAAAAGTGTCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATTGCAATTATTAAATTAAAGATAACCTATTATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTACTTCAATTGTAAGCAATGTCAGGGTGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTATCTGTATTTCTAATTCTACAATGAAGATGA
ATTCCCTGTATAAAATAAGAAAATTAAATCTGAGGTAAGCAGAGCAGACATCATCTG
TTGCTCTAGCCTCACTTCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTTAGTGTACGGAAACAGATCTCAGCAAAGCCACTGAGGAGGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTGGTAAAGGAACTAAAGAACAAAATCATCTGGTAATTCTTCTAGAAG
GATCACAGCCCCCTGGATTCCAAGGCATTGGATCCAGTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGCTCCCCCTCAAATTCACTCTTGGAAATTCTCAGTCTGTGAGTTATTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCCGGGAAGACTATGTA
AAGATGAAGGCAGAGATGGAGTTTGAGGCCACAAGCTAAGAAACACCAAGGATTGGCAACC
ATCAGAAGCTGGAAGAGGAAAGAAGATTCTCCCTAGAGGTTAGAGGATAACGGCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGAATAATTCTGGCTGTTTAA
GCCACCAAGGATAATTGGTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCGTGTATTACATTCTGTGTGCTCCCTCCACATGTACCCAAAGTTGTCTTGACCAA
TAGAATATGGCAGAAGTGTGGCATGCCACTTCAAGATTAGGTTATAAAAGACACTGCGACTTC
TACTTGAGGCCCTCTCTGCCACCCACGCCCAATCTATCTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCTATAAGAGACTTACGTGGTAAAAAAATGAAGTCTCCT
GCCACAGGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTGTTT
AAGTTGCTCAGTTGGTCTAATTGTTATGCAAGCAATAGATAAAATATGAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEEISVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTCTTCTTATGGGGACCCCTGGCCACCAGCTGCCCTCTCTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGGAGAAACTGTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAAGAAGTGCTGTTCCCTCAATCTGATAGGTC
CAGCCTTATATGCAGGAGGTGGTGCCTCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCCTTAGTTACAAAGGAAACCAATGCCACTTTGTTTATA
AGACCCAGAAGGTAGACTTCTAAGCATAGATATTATGATAACATTCACTGTAACGGTGTTC
TATACACAGAAAACAATTATTTAAATAATTGTCTTTCCATAAAAAAGATTACTTCCAT
TCCTTAGGGGAAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTTATTTATAAA
TGTATTTATTATTATAAGACTGCATTTATTATATCATTATTAATATGGATTATTTAT
AGAAACATCATCGATATTGCTACTTGAGTGTAAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKVSSFLMGTATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRILIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFQPSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDVTKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTCCCCACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAAGGTCAGTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGCTGGGGCCAAGTGGAGTGAGAAACTGGATCCCAGGGGAGGGTGCCAGAT
GAGGGAGCGACCAGATTAGGTGAGGACAGTTCTCATTAGCCTTTCTACAGGTGGTGCAT
TCTGGCAATGGTCAATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCAGTGCCTGTGCCTCCCTAGAGCCTGCTAG
GCCCAACCGCACCCAGAGCTCTGAGGAGCTGCTGAGGTGGAGCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGCTCCCCCAGGACCTGTACCAACGCCCCGT
TGCCCTGTGCCCGACTGCCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGCGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGCCCTTTGGAAAACCTGGAGGCCAGGTGTACA
ACCACTGCCATGAAGGGCAGGATGCCAGATGCTTGGCCCTGTGAAGTGCTGTGGAGCAG
CAGGATCCGGGACAGGATGGGGGCTTGGGGAAAACCTGCACCTCTGCACATTGGAAAAGAG
CAGCTGCTGCTTAGGCCCGGAAGCTGGTGTCCGTCACTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACCTGTCTTCCCTTTCCATCCCTGCTACCCCTG
GCCCAAGCAGGCACCTCTAGATATTCCCCCTGCTGGAGAAGAAAGAGCCCTGGTTTATT
TTGACATGGATGATTCTGAGGAGGAAGCTGTATTGAATGTATAGAGATTATCCAAATAAATAT
CTTATTTAAAAATGAAAAA

MRERPRGLGEDSSLISLFLQVVAFLAMVMGTHTYSHWPSCCPKGQDTSEELLRWSTVPVPPLEPA

RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGGATGTCGCTCGTGTGCTAAGCCTGGCCGCCGTGTGCAGGAGGCCGTACCCCGAGAGCC
 GACCGTTCAATGTGGCTGAAACTGGGCCATCTCCAGAGTGGATGCTAACACATGATCTAATCC
 CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGACTATTCA
 ATTTTGATGAATGTAAGCTGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
 GATTGTTGACGGGCAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
 CCTTCCAGACTCAGACCAGACCCCTCTGGTGGTAAATGGACATTTCCTACATGGCTTCCCTGTA
 GAGCTGAACACAGTCTATTCTATTGGGGCCCATAATATTCTAATGCAAATATGAATGAAGATGG
 CCCTTCCATGTCGTGAATTCACCTCACCAAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
 AGTGTGTCAAGGCCGGAAAGCCTGTGGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
 GTAGAAGTGAACCTCACACCACCTCCCCCTGGAAACAGATACTGGCTCTATCCAACACAGCAC
 TATCATCGGGTTTCTCAGGTGTTGAGCCACACCAGAAAGAAACAAACCGCAGCTTCAGTGGTGA
 TTCCAGTGACTGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTCCTACTTGTGGC
 AGCGACTGCATCCGACATAAGGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTCCCTCT
 GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTGCTGTCTGCTGGTGGCCA
 CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTCCCTT
 TCTACCACCAACTACTGCCCTTCTTAAAGGTCTGTGGTTACCCATCTGAAATATGTTCCA
 TCACACAATTGTTACTTCACTGAATTCTCAAAACCATTGCAAGAAGTGGGTGACTCTGAA
 AGTGGCAGAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCTACTCAAAGAAGGCA
 GCAGACAAAGTCGTCTTCCCTTTCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
 GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTCCCCCTGCCTTAACCTTTCTGCA
 GTGATCTAAGAAGCCAGATTCACTGCAACAAATACGTGGTGGTCACTTTAGAGAGATTGATACA
 AAAGACGATTACAATGCTCTCAGTGTGCCCCAAGTACCACTCATGAAGGATGCCACTGCTT
 CTGTGCAGAACTCTCCATGTCAAGCAGCAGGTGTCAAGCAGGAAAAAGATCACAAGCCTGCCACG
 ATGGCTGCTCCCTTGTAG

FIGURE 158

Important features of the protein:
Signal peptide:
amino acids 1-14

Transmembrane domain:
amino acids 290-309

N-glycosylation sites.
amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.
amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.
amino acids 116-122

Amidation site.
amino acids 488-492

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCGTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCGGCAGCTCGAAAATCCCCAAAGTAGGACATA
CTTTTTCAAAGCCTGAGAGTTGCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCACCACCATGTCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLLHGPAMVKYLLL SILGLAFLSEAAARKIPKVGHFFQKPESC PPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPQQETLVV
RRKHQGCSVSFQLEKVLTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTGCTGAAGTAGGAGGAGACTCAGGACTCCCAGG
ACAGAGAGTGACAAAACCTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATTG
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCTTGGGGGGGGGCAGCAC
AGGGCCTAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTCTGTCCT
TGGCACTGGGCGAAGCCCAGTGGTCTTCTCTGGAGAGGCTGTGGGCCCTCAGGACGCTACC
CACTGCTCTCGGGCCTCTCTGCCGCTCTGGGACAGTGACATACTCTGCCCTGCCCTGGGACAT
CGTCCTGCTCCGGGCCCCGTCTGGGCCCTACGACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTGGCCGTGATGGCACTGG
GAAGAGCCTGAAGATGAGGAAAGTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCTCTCCAGGCCAAGTCGTGCTCTCCAGGCCAACCTACTGCCGCTGCCCTG
TGGAGGTGCAAGTGCCTGCTGCCCTTGTGAGTTGGTCAAGTGTGCTGTGGTATATGAC
TGCTCGAGGCTGCCCTAGGGAGTGAGGATCGAACATCTGGCTTAACGTGTCAGCAGATGGTGACA
GGAACTCAACACACACAGCAGCTGCCCTGCCCTGGCTAACGTGTCAGCAGATGGTGACA
ACGTGCATCTGGTCTGAATGTCTCTGAGGAGCAGCACTCGGCCCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGA
CCACACAGACCTGGTCTGCCCTGTATTAGGTGTTGGCCTCTGGAAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCCCGCGCACACCAGAACCTCTGGCAAGCCGCGACTG
CGACTGCTGACCCCTGCAGAGCTGGCTGCTGGACGCCGACTGGCTCGCTGCCCGCAGAACGGGACT
GTGCTGGGGGCTCCGGGGGGGACCCCTGCCAGGCCACTGGTCCACCGCTTCTGGAGAACG
TCACTGTGACAAGGTTCTGAGTTCCTGAGGATCGAACATCTGGTGGGCTCTGTGTTCAAGGTG
AACAGCTGGAGAAGCTGCACTGAGGAGTGCTTGTGGGCTGACTCCCTGGGCTCTCAAAGA
CGATGTGCTACTGTTGGAGACACGAGGGCCCCCAGGACAACAGATCCCTGTGCTTGGAAACCA
GTGGCTGTACTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTGCCCTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCTGCCTACTCTTGC
CTCGCTTCCCTCATCTCCCTCTCAAAAGGATCACCGAAAGGGTGGCTGAGGCTTGTGAAA
CAGGACGTCCGCTGG
CTCGGTTTCAGGGCCCTGG
CCGTAGACCTGTGGAGCCGCTGTGAACCTGAGCGCCAGGGGGGGGGGGGGGGGGGGGGGGGG
CGGGGCCAGACCCCTGCAGGAGGGCGGGCTGGTGGCTTGTGCTCTCTCCGGTGGCTGGCGCT
GTGCACTGGCTACAGGATGGGGTGTCCGGGGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GCCGCTCGCTCAGCTGCCGCTGCCGACTCTTGCAAGGGGGGGGGGGGGGGGGGGGGGGGGGG
CACACTGCCCTCCCAACTGCCAGACTTCTGGGGGCCCTGCAGCAGCCTCGGCCGGGGGGGG
GGGGGCTCCAAGAGAGAGGGAGCAAGTGTCCCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
CATCCCCCGGGGACTCCCGCCGGGGACGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
CGGGACTTAAAAAGGAGCAGCGCTGTTTCTAAAAAAA

FIGURE 162

MPVWFLLSLALGRSPVVLSLERLVGPQDATHCSPGLSCLWDSDILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRLVAHVHLAVGHWEEPEDEEKFGGAADSGVEEPRNASLQAQVVLSPQAYPTARCVLLEV
QVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVS
EEQHFGSLSYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSSVRTNICPFREDPRAHQN
LWQAARLRLLTLQSLLDAPCSLPAEAALCWRAPIGGDPCQPLVPPLSWENVTVDKVLFFPLLKHPNLCVQ
VNSSEKLQLQECLWLADSLGPLKDDVLLLETGPQDNRSLSALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
RAAILLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVVLLFSP
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFLGALQQPRAPRSGLQERAEQVSRALQPALDSYFHPPGTAPGRGVGPAGPGAGDGT

Signal sequence:
amino acids 1-20

Transmembrane domain.
amino acids 453-475

N-glycosylation sites.
amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.
amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 552-555

N-myristoylation sites.
amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCACTCTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCAGTGAATTCCAGTCCAGCAACTTTGA
AAACATCCTGACGTGGACAGCGGGCCAGAGGGCACCCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGCAAAGAAGGGCTGTCAQCGGATCACCGGAAGCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTCACCGCT
GTCAGTGCAGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTCACTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATGATTGAGATGATTGTTATCCTA
CCCCCACGCCAATCCGTGCAGCGATGCCACCGGTAACCTGGAAGACATCTTCCATGACCTG
TTC TACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAAATGACACCTTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTGGCTGACCCCTGACACAGAGTCTCTGGCACCATCATGATTTCGTTCTCCA
CCTGGGCAAGGAGAGTGCCTCTACATGTGCCAGTGAAGACACTGCCAGACCGGACATGGACC
TACTCCTCTCCGGAGCCTCCCTGTTCTCATGGGCTCCTCGCAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGCCGCTGCACCTCCAACTCCCTGAACGTCAGCGAGTCTGACTT
TCCAGCGCTGCGCTCATCCAGGAGCACGTCCTGATCCCTGTTGACCTCAGCGGGCCCA
AGTCTGGCCCAGCCTGTCAGTACTCCAGATCAGGGTGTCTGGGACCCAGGGAGCCGAGGAGC
TCCACAGCGGATAGCCTGTCAGTACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCATTACGCCCCACA
GCCCATCTAGGTCCAGGCTTCTCTCATGGCCCTCAAGCCTCGGAGCTGGCTGGCC
CCTATGGGTATGCAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTCTAGTCCTAAA
CACCTTAGGCTAAAGGTCAAGCTCAGAAAGAGCCACAGCTGAAAGCTGCATGTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCTGGCTATGGAGGAATCCAAAGAAGCAAATCATGCA
CCCTGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCAGCTCCCCCTCTCCTCAGTCCAGATGAGGGCCACCCATGTC
CCTCCCTTGCAACCTCCTCCGGTCCATGTCCTCTGGACCAAGGTCAGTCCCTGGGGCC
TGCTGGAGTCCCTGTTGTCAGGATGAAGCCAAGAGGCCAGGGCTGGCCCTGAGACCTCAGACCTG
GAGCAGCCCCACAAGAACTGGATTCTCTGGAGGCTCAGAGGCTGGCCCTGACTGTCAGTGGGAGTCTG
AGGGAATGGGAAGGGCTGGTGCCTCCCTGTCCTCAGTGTCAATCCTGGCTGTCA
ATCCCATGCCTGCCCAGTCACACACTCTGCATCTGGCTCAGACGGGTTGCCCC
AGAGGGAGTGGCATGCAGGGCCCTGCATGGGTCCGCTCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTCTGGGAGCAGCTGTGTTAGACAAGCGCGTCTCGCTGAGCCCTG
CAAGGCAGAAATGACAGTGCAGGAGGAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCAGGATTCAAAGTGCAGGAAATTGCTCTCCCTGCCCAITCCTGGCAGTTTC
ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCTCTGTCTGTCATTGTTAAAGGTGGGAAGAGA
GCCCTGGAAAAGAACAGGCTGGAAAAAGAACAGAGAAGGGCTGGCCAGAACCCAGAACACCTGC
ACTCTGCCAAGGCCAGGGCAGCAGGAGGACTCTAGGGAGGGTGTGGCTGCAGCTCA
TTCCCAAGCCAGGGCAACTGCCTGACGTGTCACGATTTCAGCTCAATTCTCTGATAGAACAAAGC
GAAATGCAGGTTCCACCAAGGGAGGGAGACACACAAGCTTCTGCAAGGAGGAGTTCA
ATCTCTGAGAATGGGTTGGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
TGTACTGATGTCACAACCTTGCAAGCTGCTGCCTGGGTTCAAGCCATCTGGCTCAAATTCCAGC
CTCACCACTCACAAGCTGTGACTTCAAACAAATGAAATCAGTGCCTGAGAACCTCGGTTCTC
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTAAGATGAAATGAAGTCATG
TCTTAAAGTGTCTAATAGTGCCTGGTACATGGGAGTGCCTGAGTGCCTGAGTGTCTTAAAAA
AAAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVT CIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTFQPLRFIQEHLVLPVFDSLGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPILLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESLVC PKDEAKSPAPETSDLEQPTELDSLFRGLALTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCCCGGTGGCCACAACATGG
CTGGCGGCCGGGCTGCTCTCTGGCTGTCGTGGGGCGCTCTGGGGTCCGGGCCAG
TCGGATCTCAGGCACGGACGGCGTTCTGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGCGTTTGTAATTAA
AAAAAGGTGACGATGTATATGTCTACTACAAACTGGCAGGGGATCCCTGAACCTTGGCTGGA
AGTGTGAACACAGTTGGATATTTCAAAAGATTGATCAAGGTACTTCATAAAATACCGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTGTCTGTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGCGTGA
ACTTGACCCCTGTGCCTGAGCCCGAGGCATTCAGAGCTGATTCAAGGGATGGAGAAGGTGCTTCT
CAGAGAGCACCGAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGCTTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAAGTGCAGGGAAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTTCGTTGGCATAAAATCTAAGTTGTTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTGGCAGTTGCATTTGAGCAAACAAAAATATATTATTTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVC GDEEC SMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHI PADETDFVCFEGGRD
DFNSYNVEELLGSLEEDSVPEESKKAEEVSQHREKSPEESRGRELDPVPEAFRADSEDGEGA
FSESTEGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCIVIYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTTCTGCCTCTAGGACATACACGGGACCCCCCTAACTTC
AGTCCCCCAAACGCGCACCTCGAAGTCTGAACTCCAGCCCCGACATCCACGCGCGACAGG
CGCCGCAGGCAGGTCAGGCGCAGGGCGAAGGCAGTGCAGGGGCTGGGCAGCTGGCTGGC
GGCGGGAGTAGGGCCCAGGGAGGCAGGGAGGCTGCATATTCAAGTCAGCTGGCGGGCTGC
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGAGTGGCGTGG
TCTCGCTGCTGCTGGCGCGCTGCTGCGGCCACGGAGCCTCTGCGCCGCGTGGTCAGC
GGCCAAAAGGTGTGTTCTGACTTCAGGAGGCAGGCCGCTGGCTGTGAGAGTGAGGGAGGAGT
GTCCAGCCGAGTGAGCTTCAGGAGGCAGGCCGCTGGCTGTGAGAGTGAGGGAGGAGT
GCCTTGAGAATGAAGCAGAACAGAAGTAATAGAGAGCATGTTGCAAAACCTGACAAA
ACAGGGATTCGATGGTATTCGATAGGGCTTGGAGGAATGGAGATGGCAAACATCTGG
TGCTGCCAGATCTTACCACTGGCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGATGTATACCAACCAACTGCCAATCCTGGC
CTTGGGGTCCCTACCTTACAGTGGATGATGACAGGTGTAACATGAAGCACAATTATTTG
CAAGTATGAACCAGAGATAATCCAACAGCCCTGAGAAAAGCCTATCTAACAAATCAACCAG
GAGACACCCATCAGAATGGTTACTGAGCAGGTATAATTCCCAACTTAATTATGGT
ATACCAACAATACCCCTGCTTACTGATACTGGTGTGCTTGGAAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAGGAAGAACAAAAGTACCTGAAACTCTAACACTGTGATTCAAAGAGTA
CCAGAAAAGAAGTGGCATGGAAAGTATAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTGGAAATCACAAGGATCTGC
AAGATGAACTGTAAGCTCCCCTGAGGCAAATATTAAAGTAATTGTTATGTCTATTATTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCC
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAGTTGTTATCAACACGT
GGTGTGTTAGAGCAATTCTTTATTCTTACCTTCTACAGTGTGTTACTGCAATGTAA
TGTATATTGTATTGAAATTACAGTGTGCAAAGTATTTCACCTTGTGCTAAAGTGTGTT
ATGAACCTGTTCAATATTATTATTGTCATCTACATTTCACATGCTTGTGATTAAAG
AAACTTATTACTGTTGTCAGTAATTGCAACACACACAAATATAGTACCATAGAAA
TTTCTGAAATAATTCTACATTGCTCAGCTCTGCTTGGTCAATGTCTAGGAAATCTTCAGA
AATAAGAAGCTATTCTACATTAGTGTGATATAACCTCCTCAACACATTAACTTAG
TGTCTAATTCTACATTGTGCAAGACATGTGCTTATAATTATTCTTAAACAGATT
TTGTAATAATGTAACTTTGTAATAGGTGCTAAACACTAATGCACTGCAATTGAA
TGACATACACAATATAAAATCATATGTCCTCACACGTTGCCTATATAATGAGAAG
GGGTTCTGAAATCAATGTGTCCTCTTGCCCCTAAACAAAGATGGTTGTTGGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATT
ACTATATTAGTATACAAAGAGGTCTGTTGAGACCAGGTGAATAGTC
ACAAGCACAGCACACAGACATTAGGAAGGAAAGGAAC
ACAGAAATCGTGTGAAAATGGTTGG
AACCCATCAGTGATGCGATATTCTACATTGATGAGGGTTGCTTGA
GAGATAGAAAATGGTGGCTCTT
CTGCTTATCTCTAGTTCTCAATGCTTACGCCCTGTTCTCAAGAG
AAAGTTGTAACTCT
CTGGTCTCATATGTCCTGCTCTTAAACCAAA
AAAAAGAGTTCTGGGGAAAA
AAAAAAAAAAAAAAAAAAAAA
A

FIGURE 168

MSRVSLLGAAALLCGHAFCRVVSGQKVCFADFKHPCYKMAKFHELSSRVSFQEARIACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217